STEREOLOGY -

A STATISTICAL VIEWPOINT

by

Pamela Joy Davy

A thesis submitted for the degree of

Doctor of Philosophy

of the

Australian National University

March, 1978

PREFACE

The theory of stereology exploits results from branches of mathematics such as integral geometry with which statisticians are not generally familiar, and thus a certain amount of survey material has been included to render the thesis readable to a reasonably broad audience. The sources of such material have been quoted wherever it appears in the text.

The remainder was either my own original work or was done jointly with R.E. Miles. Several prior publications form a basis for the thesis, namely Davy (1976, 1977), Davy and Miles (1977) and Miles and Davy (1976a, 1976b, 1977, 1978). It is difficult to divide the joint work into contributions made by the respective authors, as it evolved through continued interchange of ideas, but an approximate breakdown is as follows.

R.E.M. realised that integral geometry could be used to derive certain stereological estimators; I extended such estimation to centroids, integral-geometric measures and vector fields, and worked out the variances in Chapter 4. In order to eliminate the bias which R.E.M. realised to exist for ratio estimators, I suggested the use of A-weighted planar sections, worked out how they could be generated, and made comparisons of mean square error. R.E.M. generalised this idea to n dimensions, and pointed out that weighting could be applied to more general types of probe such as grids and quadrats. R.E.M. worked out thick section formulae for a 3-dimensional Poisson grain process - I extended these to n dimensions and to cylinders as well as grains. R.E.M. suggested the use of wedge probes to estimate Gaussian curvature; I obtained the quadratic equation relating the principal curvatures at the intersection of a wedge probe with a surface, and found a conditional orientation argument which enabled estimation of $\kappa_1 \kappa_2$.

Pamela Davy

ACKNOWLEDGEMENTS

As is usual with research, this thesis did not arise out of work in isolation but rather was nurtured by discussions with many people, to all of whom I wish to express my gratitude even if they are not specifically mentioned below.

In particular, I thank my supervisor, Dr R.E. Miles, for his constant support, both moral and mathematical, his patience, his imagination and his insight. As stereology has many facets, my collaborators were varied in their interests. On the probabilistic side were Professor P.A.P. Moran and Mr A. Baddeley. On the computational side were Dr R.S. Anderssen and Dr A. J. Jakeman. And on the practical side were Professor E.R. Weibel and Dr H. Keller of Berne. I spent two months in Berne at the end of 1975 with the help of a travel grant from the Roche Foundation, and here obtained most of the motivation for my thesis.

I thank Mrs B. Geary who typed the thesis and Mrs B. Cranston who typed the papers which formed the basis for it - their perseverance with my notation is to be commended.

The environment of the Australian National University was one which suited me well. I was supported during my three years of study here by a Commonwealth Postgraduate Research Award.

I am indebted to Graeme Chandler for his assistance during the long hours of proof-reading.

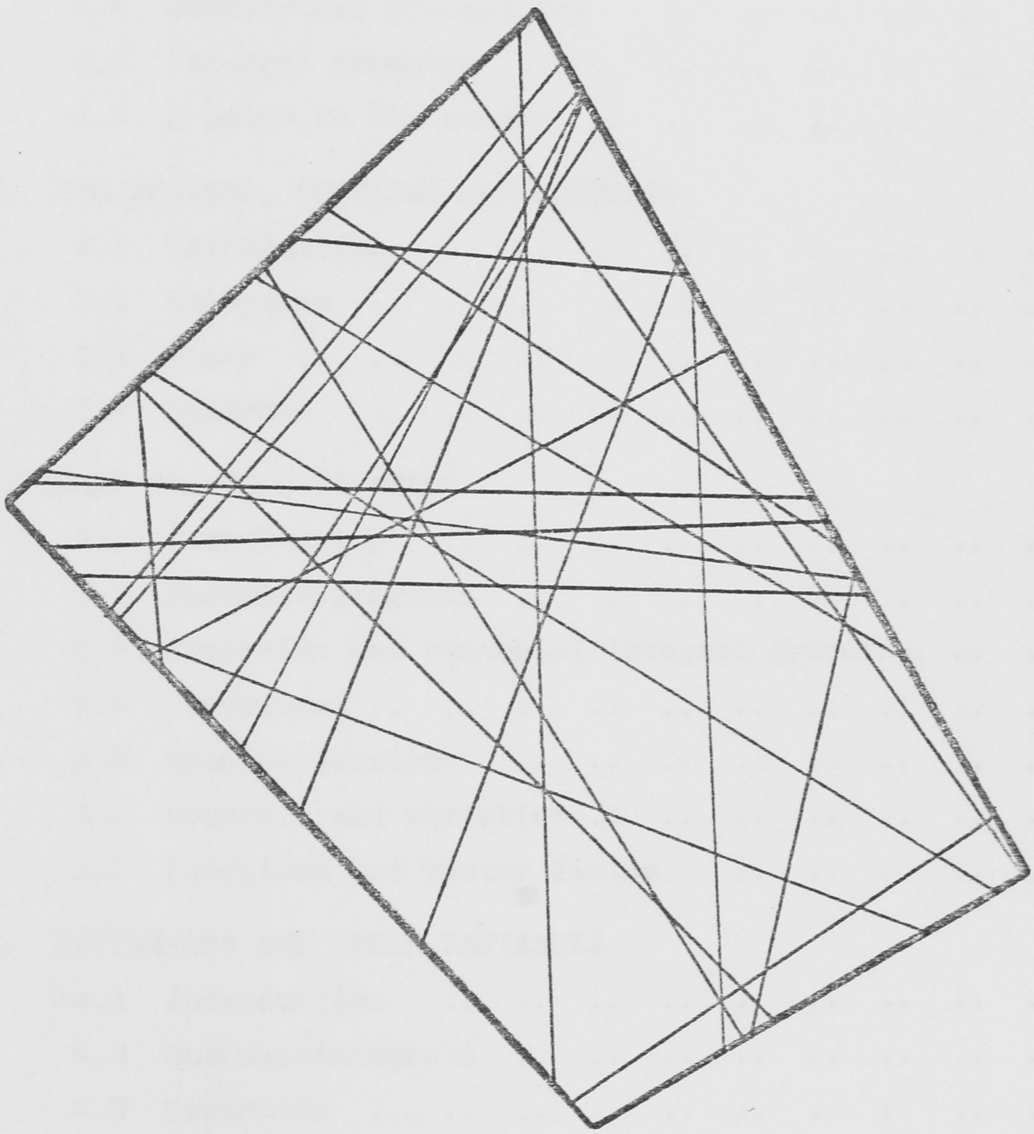
ABSTRACT

Various geometrical properties of a feature set contained within a compact n-dimensional specimen X may be inferred by generating a random position of a probe intersecting X or by projecting X in an isotropic random direction. The variance associated with such sampling is highly structure dependent and may be expressed in terms of a double integral over X. Ratio estimation, in the form of a set of fundamental formulae, is widely used in stereology, and is shown to be unbiased with the use of appropriately weighted probes, which are relatively easy to generate in a geometrical setting. Certain advantages other than elimination of bias are inherent in weighted sampling - in particular the mean square error is, in certain circumstances, reduced. Besides providing a rigorous derivation (which, except under very restricted conditions, has been lacking even for 2 or 3 dimensions), this thesis provides new stereological formulae of potential application. Variants of standard sectioning, namely wedge sections and curved sections, are also considered, with the interesting outcome that some properties not accessible by flat sectioning are thereby estimable.

Another interpretation of stereology is given in terms of random sets of infinite extent. The fundamental formulae in this case become identities between certain densities, which can be defined both as expected values of geometrical quantities per unit volume and as coefficients associated with the intersection of a fixed compact convex set with the random set.

One practical problem in transmission microscopy is that of correcting the fundamental formulae when a slice of non-zero thickness is observed instead of a true section. A multi-dimensional treatment of this problem is given. Multiple sampling, either systematic or independent, is important to increase the precision of stereological estimates. Multistage

procedures, which can be formulated in terms of martingales, are also investigated. Certain techniques of Monte Carlo estimation have a special interpretation in the context of stereology.



FRONTISPIECE. 24 independent *L*-weighted lines through a planar specimen

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION TO STEREOLOGY

1.1. History and development

The word "stereology" is of recent origin, having been chosen in 1961 at a meeting of 11 scientists who were interested in obtaining 3-dimensional information from 2-dimensional sections or projections. It was realised that certain problems which had arisen in diverse subject areas were in fact equivalent and could be solved by the joint use of probability and geometry. A definition of stereology as the extrapolation from 2- to 3- dimensional space has been suggested by Elias (1967) - however to embrace the use of point and line data, a better definition would be extrapolation from lower to higher dimensional space. Sometimes the restriction of data to a lower dimension is enforced by available measurement procedures; at other times it enables ease and speed of measurement when confronted with a surfeit of information. Morphology, morphometry, quantitative microscopy, quantitative metallography and stereometry are disciplines which overlap with stereology, and whose names appear frequently in the literature, especially prior to 1961.

In spite of the youth of its name, stereology has its roots in the nineteenth century. In 1847, Delesse realised that in order to determine the relative proportions by volume of minerals within a piece of rock, it is not necessary to resort to crushing and separation. Rather, a planar section of the rock can be polished and the relative areas of the different phases measured. In 1898, Rosiwal carried the dimensional reduction one step further: he determined volume fractions by measuring the relative lengths of the phases on a set of line segments placed on a planar section. And in 1933, Glagoleff suggested an even simpler procedure - one could simply use the proportions of points occupied by the different phases when a lattice is laid out over a planar section.

The discovery of methods of estimating properties other than volume or area through lower dimensional samples, being less intuitive, came somewhat later. In 1945 Tomkeieff showed how to estimate surface area per unit volume by measuring the numbers of intersections along line segments placed through the specimen. As recently as 1967, Cahn and DeHoff independently derived the result that the integral mean curvature of a surface per unit volume can be estimated by the total curvature on a planar section.

Thus, in a timespan of over a century, a set of basic principles, called the "fundamental formulae of stereology" has evolved. Weibel (1973) gives a more detailed account of this development. In order to give the reader some appreciation of the significance of these formulae, let us look briefly at some of their fields of application.

1.2. Applications

Stereology provides objective, quantitative methods for sciences which were once primarily qualitative. This is important as the human eye can be deceived, or fail to detect subtle differences. At the same time, stereology provides interpretations of what can be seen from a limited 0-, 1- or 2-dimensional viewpoint. For an overview of stereological applications, the reader is referred to the proceedings of the four international congresses for stereology which have been held to date (Haug and Elias (eds.) (1963), Elias (ed.) (1967), Weibel *et al.* (eds.) (1972) and Underwood *et al.* (eds.) (1976)).

Life sciences which make use of stereology include anatomy, botany, cell biology, embryology, forestry and pharacology. Volume fraction has been used in studies of cancer development, muscle contraction and effects of drugs. Surface area is important in the diffusion capacity of lungs and the evolution of the cerebral cortex. Joint volume and surface measurements have been used, for example, in developmental studies of seeds and embryos.

Materials and earth sciences using stereology include geology, metallurgy and petrology. Volume fraction is used for quality control of steel and classification of rocks. Specific surface is related to chemical reaction sites, obstacles to dislocations, grain growth and brittleness. Integral mean curvature is related to chemical potential and pressure difference. Special materials investigated by stereological means include coal, concrete and snow.

1.3. A statistical viewpoint

Why should a statistician choose to write a thesis on stereology? Firstly, stereology is concerned with inference on the basis of a sample taken from an infinite geometrical population, and may therefore be regarded as a sub-discipline of statistics, even though the majority of mathematicians have paid no attention to it in the past. Exceptions include Moran (1972), who puts forward a probabilistic viewpoint of stereology, and Wicksell (1925, 1926), Nicholson (1970), Jakeman and Anderssen (1975) and Anderssen and Jakeman (1975), who investigate inference concerning the size distribution of a population of randomly dispersed particles in space on the basis of planar sections or linear probes.

Secondly, although the fundamental formulae of stereology have been known for a long time, they are *not* valid with the generality assumed in much of the literature. The proofs given in the standard texts of the subject (e.g. DeHoff and Rhines (eds.) (1968) and Underwood (1970)), while they are commendable in that they are accessible to non-mathematical scientists, are based on very restrictive assumptions.

For example, the Delesse equality between volume fraction $\left(V_V\right)$ and areal fraction $\left(A_A\right)$ of a certain phase in a material is generally supposed to hold true for an arbitrarily shaped specimen sliced by a plane which is random in some, usually unspecified, sense. The usual derivations of this relationship consider a cubic specimen sliced parallel to a square face, the distance from that face being uniformly distributed. This is to ensure that the area of section, represented by the subscript in A_A , is not random, thus avoiding the necessity of considering a ratio of dependent random variables. Derivations of the formula for specific surface area (S_V) involve an assumption either of the internal structure under examination or of the sampling scheme, even though the former assumption is usually false for given specimens and the latter is incompatible with a cubic specimen!

Since Nature does not always manifest herself in cubic or isotropic form, this thesis provides a set of precise, yet quite general, conditions under which these fundamental formulae, and their generalizations to n-dimensional Euclidean space, are valid. The conditions involve the sampling scheme, the structure being supposed deterministic. The main reasons for adopting an n-dimensional framework are that it compactly subsumes a large number of different practical cases and that it reveals similarities in form between seemingly disparate relationships. Besides the rigorous validation of the classical stereological estimators, many new estimators are presented. Some variances are also derived, although these tend to be highly structure dependent.

There is a kind of duality between the specimen and the sample - instead of considering a fixed specimen and random sampling scheme, we can consider a random specimen and a fixed sample. In Chapter 7 a different interpretation of the fundamental formulae is given in terms of random sets (see Davy (1977) for an elementary exposition of random sets).

The theory of stereology as presented here depends heavily upon two disciplines - geometrical probability and integral geometry. We shall now give a brief account of these and explain why they are so important.

1.4. Geometrical probability

Geometrical probability, perhaps because of its aesthetic appeal and the ease with which certain problems within it may be posed, has quite a long history. Early workers include Buffon (1777), who investigated the probability that a needle thrown onto a parallel grid of lines should intersect one of the lines, and Crofton (1885), who derived a number of elegant results concerning random points in and lines through a planar convex set.

The subject is concerned with probabilities and expected values associated with randomly located geometrical objects (e.g. points, lines, circles, planes) in space (usually Euclidean, although Santaló (1976) has worked in spaces of constant curvature). Kendall and Moran (1963) give a good introduction and an extensive bibliography, which has been updated by four further papers on recent research in geometrical probability (Moran (1966b, 1969), Little (1974) and Baddeley (1977)). In the context of stereology, the randomly located geometrical objects are the probes (i.e. sections, transects, points) whose interactions with the underlying structure are to be observed.

1.5. Integral geometry

Just as classical probability theory relies on Lebesgue measure theory, geometrical probability makes use of certain measures defined on spaces of geometrical objects. The evaluation of integrals involving such measures is known as integral geometry.

Much of the literature of this subject is in German (e.g. Blaschke (1949), Hadwiger (1957)), although two recently published books (Matheron

(1975), Santaló (1976)) include expositions in English. For convenience, many integral geometric identities are expressed in terms of differential forms. This practice has been adopted in the following chapters. See Flanders (1963) for the definition, manipulation and applications of differential forms. Convex sets play an important part in integral geometry. It turns out that certain integrals performed over the objects hitting a compact convex set can be related both to its global properties such as volume and surface area and to its local properties such as mean or Gaussian curvature.

1.6. A guide to the reader

A few brief comments are given here to aid the reading of this thesis. The two focal points are Chapters 5 and 7, which derive the fundamental formulae of stereology for random sampling (deterministic structure) and random structure respectively. Two sets of data are analysed; §10.5 corresponds to the deterministic formulation and §8.4 to the random set formulation. Chapters 2-4 set up the technical apparatus needed later, Chapter 4 containing some results which, although used to evaluate certain variances, are of pure integral geometric interest. Chapter 6 deals with variants of stereological sectioning which are probably of theoretical rather than practical interest. Chapters 8-10 deal with practical problems of stereology - sections of non-zero thickness, use of multiple observations or sampling stages, and Monte Carlo Techniques.

There is one important area of stereology which has been omitted, although I have dabbled in it. This is the estimation of distribution functions associated with a system of particles, membranes or pores on the basis of lower dimensional information. To do it justice would have required considerable lengthening of the thesis.

A glossary of notation for symbols of constant meaning is given at the end of the thesis - the reader is advised to read the comments therein regarding conventions of usage before proceeding further.

CHAPTER 2

PROJECTIONS, SECTIONS AND QUADRATS

2.1. Introduction

In classical sampling theory it is required to estimate properties of a finite population on the basis of a random sample. In stereology we have a somewhat similar situation. The finite population is replaced by a compact subset X (the *specimen*) of E^n and it is required to estimate properties of X on the basis of a randomly selected subset, possibly deformed by the observation process, and usually of dimension lower than n. The notion of a simple random sample of a finite population is straightforward, while that of a random subset requires further clarification. This chapter explores methods of obtaining samples from an n-dimensional specimen.

The simplest example is the selection of a random point of X. Historically, geometrical probabilists have used "random point" to denote a point which is equally likely to occur anywhere within X. Formally, this point x has probability element $d\mathrm{x}/V$, where $d\mathrm{x}$ refers to n-dimensional Lebesgue measure and V is the n-volume of X, assumed here to be non-zero. When we turn to subsets more complicated than points, however, the interpretation of randomness is less obvious. As illustrated by Kendall and Moran (1963), the phrase "random chord of a circle" is ambiguous. Intuition demands some kind of invariance but it is not clear what form this should take.

Lebesgue measure seems natural to use in connection with random points due to its invariance under translation. In general, Haar measures on topological homogeneous spaces acted upon by groups of Euclidean motions are useful tools for constructing random sampling schemes on X. Nachbin

(1965) gives the definition and details of existence and uniqueness of Haar measures. The next three sections deal with Haar measures for subspaces, flats and mobile sets, which enable the specification of random projections, sections and quadrats respectively.

2.2. Subspaces

Consider the space of all r-subspaces L_r acted upon by the group of rotations of E^n. This space is called the Grassman manifold. An explicit expression for the Haar measure is given by Miles (1971a). We shall need, however, only certain total integrals and identities between differential elements. The element of Haar measure will be denoted by dL_r (or dL_r^n when we wish to emphasize that we are considering subspaces of E^n), and is scaled so that the total integral

$$\int dL_r = \sigma_{n-r+1}\cdots\sigma_{n-1}\sigma_n \,/\, \sigma_1\sigma_2\cdots\sigma_r \quad (0 < r < n) \,, \tag{2.1}$$

where $\sigma_i = 2\Pi^{i/2}/\Gamma(i/2)$ is the surface area of the unit sphere in E^i. As there is a one-to-one correspondence between an r-subspace L_r and its orthogonal complement L_{n-r}, and by the choice of scaling in (2.1),

$$dL_r = dL_{n-r} \quad (0 < r < n) \,. \tag{2.2}$$

Suppose now that we restrict attention to the subspaces $L_{r(q)}$ containing a fixed q-subspace L_q $(q < r)$, acted upon by the rotations which leave L_q invariant. It is intuitively obvious, and may be shown rigorously, that the Haar measure in this case satisfies

$$dL_{r(q)} = dL_{r-q}^{n-q} \quad (0 < q < r < n) \,. \tag{2.3}$$

where L_{r-q}^{n-q} is the intersection of $L_{r(q)}$ with the orthogonal complement of L_q , and dL_{r-q}^{n-q} corresponds to Haar measure within E^{n-q} .

It will be convenient to extend the above notation in the following ways. L_0, L_n denote the origin and the entire space E^n respectively. $L_{r(r)}$ denotes a fixed r-subspace and $L_{r(0)}$ is used interchangeably with L_r . The elements dL_0, dL_n and $dL_{r(r)}$ are to be interpreted as assigning unit mass to L_0, L_n and $L_{r(r)}$ respectively. With these conventions, equations (2.2) and (2.3) remain valid in the range $(0 \leq q \leq r \leq n)$.

Consider the joint element associated with an r-subspace L_r and a q-subspace L_q^r of L_r . It can be shown (Santaló (1955)) that

$$dL_q^r dL_r = dL_{r(q)} dL_q \quad (0 \leq q \leq r \leq n) . \tag{2.4}$$

The above Haar measures may be normalized to give probability measures on subspaces. An *isotropic random* (IR) r-subspace has probability element

$$dL_r / \int dL_r \quad (0 < r < n) . \tag{2.5}$$

As an example, in the case $n = 3$, $r = 2$,

$$dL_2 = \sin\theta d\theta d\varphi / 2\pi \quad (0 \leq \theta \leq \pi/2,\ 0 \leq \varphi < 2\pi) \tag{2.6}$$

is the element for an IR plane containing the origin.

There is no immediately obvious correspondence between subspaces of E^n and subsets of X - we shall take $Z(L_r, X)$ to consist of those points which can be "seen" from L_r , i.e.

$$Z(L_r, X) = (L_r \cap X) \cup \Big\{ x \in X \cap L_r^C \mid \text{the orthogonal line segment from } x \text{ to } L_r \text{ contains no points of } X \text{ other than } x \Big\} .$$

Usually what is observed rather than $Z(L_r, X)$ is its orthogonally projected image onto L_r, denoted by $X|L_r$. Thus equation (2.5) can be used to define an isotropic random projection of X.

2.3. Flats

An r-flat F_r is a translate of an r-subspace. (A 0-flat is just a point.) It can be specified by the ordered pair (L_r, x_{n-r}), where L_r is the parallel subspace and x_{n-r} is the intersection of F_r with the orthogonal subspace L_{n-r}. When we fix the orientation L_r, the Haar measure under the group of translations of E^n is simply $(n-r)$-dimensional Lebesgue measure:

$$dF_{\underline{r}} = d\mathrm{x}_{n-r} \quad (0 \le r < n) . \tag{2.7}$$

(The bar is used to distinguish this from the variable orientation case.)

For the space of all r-flats, the element of Haar measure corresponding to the group of Euclidean motions is the product

$$dF_r = d\mathrm{x}_{n-r} dL_r \quad (0 \le r < n) . \tag{2.8}$$

An analogue of (2.4) exists for flats $F_{r(q)}$ which contain a fixed q-flat F_q, viz.

$$dF_q^r dF_r = dF_{r(q)} dF_q \quad (0 \le q < r < n) , \tag{2.9}$$

where $dF_{r(q)} = dL_{r(q)}$.

Equations (2.7) and (2.8) give rise to divergent integrals when integrated over their entire ranges, but if we consider only those flats which hit (i.e. have non-void intersection with, denoted by "$\uparrow$") X, then

$$\int_{F_{\underline{r}} \uparrow X} dF_{\underline{r}} = \int_{X|L_{n-r}} d\mathrm{x}_{n-r} = V_{n-r}(X|L_{n-r}) , \tag{2.10}$$

(V_{n-r} denotes $(n-r)$-volume; as X is compact its orthogonal projection $X|L_{n-r}$ is also compact and therefore measurable) and

$$\int_{F_r \uparrow X} dF_r = \int V_{n-r}(X|L_{n-r})\,dL_r = \int V_{n-r}(X|L_{n-r})\,dL_{n-r}$$
$$= \int dL_{n-r}.E(V_{n-r})$$

(where expectation is with respect to an IR $(n-r)$-subspace)

$$= \int dL_r.M_{n-r}(X) \text{ , say.} \tag{2.11}$$

If the right-hand sides of (2.10) and (2.11) are non-zero (which is the case, for example, if X has non-zero n-volume), then we may define a *fixed orientation uniform random* (FUR) r-flat through X and an *isotropic uniform random* (IUR) r-flat through X as having probability elements $d\underline{F}_r/V_{n-r}(X|L_{n-r})$ and $dF_r/\int dL_r.M_{n-r}(X)$ respectively. Such r-flats determine random cross-sections $X \cap \underline{F}_r$ and $X \cap F_r$ of X.

There are other ways in which we can construct random r-flats through X. Suppose that we first generate a FUR (resp. IUR) q-flat through X and then select the FUR r-flat (resp. an independent IR r-flat) containing this q-flat. From (2.7)-(2.11), the joint probability element is

FUR case	IUR case
$\dfrac{d\mathrm{x}_{n-q}}{V_{n-q}(X\|L_{n-q})} = \dfrac{d\mathrm{x}_{r-q}\,d\mathrm{x}_{n-r}}{V_{n-q}}$ $(L_q \subset L_r)$	$\dfrac{dF_q}{\int dL_q M_{n-q}(X)} \cdot \dfrac{dF_{r(q)}}{\int dL_{r(q)}} = \dfrac{dF_q^r\,dF_r}{\int dL_q \int dL_{r(q)} M_{n-q}}$

$$(0 \le q < r < n) \text{ .} \tag{2.12}$$

Integrating out the intermediary q-flat, we obtain the marginal probability element

$$\frac{V_{r-q}(X\cap\underline{F}_r|L^r_{r-q})}{V_{n-q}}\,d\underline{F}_r \quad \Bigg| \quad \frac{M^r_{r-q}(X\cap F_r)}{M_{n-q}.\int dL_r}\,dF_r \quad \text{(using (2.4)).} \tag{2.13}$$

The interpretation of (2.13) is as follows. Generating an r-flat in the above fashion is equivalent to generating a FUR (resp. IUR) r-flat through X *weighted* according to the projection $V_{r-q}\left[(X \cap \underline{F}_r) | L^r_{r-q}\right]$ (resp. mean projection $M^r_{r-q}(X \cap F_r)$) of the resulting cross-section. The terminology "V_{r-q}- (resp. M^r_{r-q}-) weighted" will be adopted.

As (2.13) contains two probability elements whose total integrals must equal 1 , we obtain the identities

$$\int V_{r-q}\left[(X \cap \underline{F}_r) | L^r_{r-q}\right] d\underline{F}_r = V_{n-q}(X | L_{n-q}) \tag{2.14}$$

and

$$\int M^r_{r-q}(X \cap F_r) dF_r = \int dL_r \cdot M_{n-q}(X) \ . \tag{2.15}$$

We may now ask what happens if a V_{q-t}-weighted (resp. M^q_{q-t}-weighted) q-flat is used in the above construction. The joint probability element (2.12) becomes

$$V_{q-t}\left[X \cap \underline{F}_q | L^q_{q-t}\right] \cdot d\underline{F}_q =$$

$$= \frac{V_{q-t}\left[(X \cap \underline{F}^r_q) | L^q_{q-t}\right]}{V_{r-t}\left[(X \cap \underline{F}_r) | L^r_{r-t}\right]} \cdot \frac{V_{r-t}\left[(X \cap \underline{F}_r) | L^r_{r-t}\right]}{V_{n-t}} d\underline{F}^r_q d\underline{F}_r \tag{2.16}$$

$$(L_t \subset L_q \subset L_r) \ .$$

in the FUR case and

$$\frac{M^q_{q-t}(X \cap F_q)}{\int dL_q M_{n-t} \int dL_{r(q)}} dF_{r(q)} dF_q$$

$$= \frac{M^q_{q-t}(X \cap F^r_q)}{\int dL^r_q \cdot M^r_{r-t}(X \cap F_r)} \cdot \frac{M^r_{r-t}(X \cap F_r)}{\int dL_r \cdot M_{n-t}} dF^r_q dF_r \quad \text{(using (2.4))}. \tag{2.17}$$

Integrating out $\underline{F}_q$ (resp. F_q) we obtain the marginal element (with the aid of (2.14) and (2.15))

$$\frac{V_{r-t}\left(X\cap \underline{F}_r \,|\, L^r_{r-t}\right)}{V_{n-t}} \cdot d\underline{F}_r \;\Bigg|\; \frac{M^r_{r-t}\left(X\cap F_r\right)}{\int dL_r . M_{n-t}}\, dF_r \quad (0 \le t \le q < r < n) \;, \tag{2.18}$$

i.e. that of a V_{r-t}- (resp. M^r_{r-t}-) weighted r-flat.

The interpretation of these results is aided by a diagram:

FUR CASE:

$$\begin{tikzcd} X \arrow[r, "V_{r-t}\text{-weighting}"] \arrow[d, "V_{q-t}\text{-weighting}"'] & \underline{F}_r \uparrow X \arrow[d, "V_{q-t}\text{-weighting}"] \\ \underline{F}_q \uparrow X \arrow[r] & \underline{F}_q \subset \underline{F}_r \uparrow X \end{tikzcd} \tag{2.19}$$

(i) A V_{q-t}-weighted q-flat through the cross-section corresponding to a V_{r-t}-weighted r-flat through X is stochastically equivalent to a V_{q-t}-weighted q-flat through X.

(ii) The r-flat containing a V_{q-t}-weighted q-flat through X is equivalent to a V_{r-t}-weighted r-flat through X.

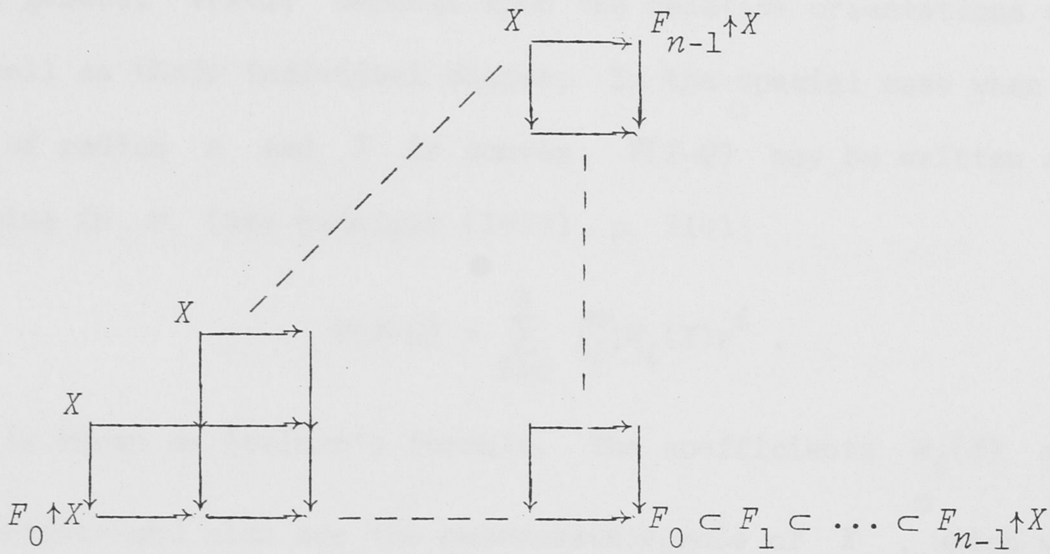
IUR CASE:

$$\begin{tikzcd} X \arrow[r, "M^r_{r-t}\text{-weighting}"] \arrow[d, "M^q_{q-t}\text{-weighting}"'] & F_r \uparrow X \arrow[d, "M^q_{q-t}\text{-weighting}"] \\ F_q \uparrow X \arrow[r, "\text{IR generation}"'] & F_q \subset F_r \uparrow X \end{tikzcd} \tag{2.20}$$

(i) A M^q_{q-t}-weighted q-flat through the cross-section corresponding to a M^r_{r-t}-weighted r-flat through X is stochastically equivalent to a M^q_{q-t}-weighted q-flat through X.

(ii) An IR $\tilde{r}$-flat containing a M^q_{q-t}-weighted q-flat through X is equivalent to a M^r_{r-t}-weighted r-flat through X.

Practical examples of these results will be given in later chapters. Note that (2.19) and (2.20) may be iterated to yield commutative diagrams of the following form:



(2.21)

2.4. Quadrats

Instead of random flats hitting X , this section deals with a rigid compact set Q (called a quadrat) whose location and possibly orientation are to be randomized over the positions hitting X .

Let Q_{y} denote the translate of Q by the vector y . A fixed orientation uniform random (FUR) quadrat Q_{y} hitting X has probability element proportional to $d\mathsf{y}$. The proportionality factor is the reciprocal of

$$\int_{Q_{\mathsf{y}}\uparrow X} d\mathsf{y} \ . \tag{2.22}$$

To evaluate (2.22), we shall use the notion of Minkowski set addition. Put $X \pm Y = \{\mathsf{x}\pm\mathsf{y} \mid \mathsf{x} \in X, \mathsf{y} \in Y\}$ for subsets X, Y of E^n .

Now, $Q_{\mathsf{y}} \uparrow X$ iff $\mathsf{q}+\mathsf{y} \in X$ for some $\mathsf{q} \in Q$

iff $\mathsf{y} = \mathsf{x} - \mathsf{q}$ for some $\mathsf{q} \in Q, \mathsf{x} \in X$

iff $\mathsf{y} \in X-Q$.

Hence $\int_{Q_y \uparrow X} dy = V(X-Q)$, which must be non-zero in order for a FUR quadrat to exist.

In general $V(X-Q)$ depends upon the relative orientations of X and Q as well as their individual shapes. In the special case when Q is an n-ball of radius r and X is convex, $V(X-Q)$ may be written as a polynomial in r (see Hadwiger (1957), p. 214):

$$V(X-Q) = \sum_{i=0}^{n} \binom{n}{i} W_i(X) r^i \ . \tag{2.23}$$

(2.23) is known as Steiner's formula. The coefficients $W_i(X)$ appearing in the right-hand side are the *quermassintegrals* of X , which will be defined in Chapter 3. For $n = 3$, (2.23) reduces to

$$V(X-Q) = V(X) + S(X)r + 2\pi M(X)r^2 + \tfrac{4}{3}\pi r^3 \ , \tag{2.24}$$

where $S(X)$, $M(X)$ are the surface area and mean caliper diameter of X respectively.

Let us suppose that the quadrat Q is a flat r-dimensional set of non-zero r-volume $V_r(Q)$. To specify location *and* orientation we need the position y of a fixed reference point on Q and orientation B_r of a fixed r-dimensional reference frame attached to Q . The element of Haar measure under Euclidean motions is

$$dQ = dy dB_r \ , \tag{2.25}$$

where dB_r corresponds to the rotation invariant measure on r-frames in E^n . An explicit expression for dB_q in terms of Pfaffion forms can be found in Miles (1971a), together with the equations

$$\int dB_r = \sigma_{n-r+1} \cdots \sigma_n \tag{2.26}$$

and

$$dB_r = dB_r^r dL_r \quad (0 < r < n) \ , \tag{2.27}$$

dB_r^r being the invariant element for r-frames within L_r .

An isotropic uniform random (IUR) quadrat hitting X has probability element

$$dy dB_r / \int V(X-Q) dB_r = dQ / \int dB_r . E[V(X-Q)] \ , \tag{2.28}$$

expectation being with respect to an IR orientation of Q . For X and Q both convex, there exists a generalization of (2.23), namely

$$E[V(X-Q)] = \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} W_i(X) W_{n-i}(Q) \ , \tag{2.29}$$

where $\omega_n = \pi^{\frac{1}{2}n}/\Gamma(\frac{1}{2}n+1)$ is the volume of the unit n-ball and W_i is the ith quermassintegral.

An alternative procedure for generating a random quadrat hitting X is as follows:

(i) independently choose uniform random (UR) points x, q of X and Q ,

(ii) locate Q so that q is superimposed onto x . In the FUR case, we are finished; in the IUR case, Q is given an IR orientation about q .

The joint probability element is

FUR case	IUR case
$dx dq / V(X) V_r(Q)$	$dx dq dB_r / V(X) V_r(Q) \int dB_r$
$= dy dq / V(X) V_r(Q)$	$= dy dq dB_r / V(X) V_r(Q) \int dB_r$, (2.30)

due to the fact that the transformation $(x, q) \mapsto (y, q)$ has Jacobian equal to 1 .

Integrating out q , we find that the marginal element for the position of Q is

$$V_r(X \cap Q) dy / V(X) V_r(Q) \quad \Big| \quad V_r(X \cap Q) dQ / V(X) V_r(Q) \int dB_r \ . \tag{2.31}$$

In other words, the invariant distribution of Q has been weighted according to the r-content of the intersection set $X \cap Q$.

To each type of random quadrat described above, there corresponds a random subset $X \cap Q$.

CHAPTER 3

WHAT CAN BE ESTIMATED?

3.1. Introduction

In the previous chapter we discussed methods of generating a random subset of the specimen set X . Before dealing with the statistics associated with such random sampling, we need to consider which properties are to be the goals of estimation. Usually in stereology the specimen consists of a two phase material. In other words, X can be partitioned as $Y \cup \left(X \cap Y^{c}\right)$, Y being a subset of X called the *feature set.* We suppose that Y is a closed subset of X whose geometrical properties are unknown. Sometimes it is desired to infer properties of X itself by sampling techniques rather than by complete measurement. The final section deals with the case where a function f is defined over X and we wish to infer some of its global properties. Of course, by taking f to be the indicator function of the feature Y , we see that this is in fact a generalization of the earlier case.

3.2. Quermassintegrals

Let C be the set of all compact convex sets. In §2.4, the quermass-integral functionals were mentioned briefly. These functionals may be defined in a number of ways, but perhaps the simplest is to put

$$W_i(Y) = \omega_n \int_{F_i \uparrow Y} dF_i/\omega_{n-i} \int dL_i \quad (0 \le i \le n,\ Y \in C) . \tag{3.1}$$

(Recall that ω_i is the i-volume of the unit i-ball.) Thus $W_i(Y)$ is proportional to the total measure of i-flats passing through Y , the constant of proportionality having been chosen in such a way that when Y

is the unit n-ball, $W_i(Y) = \omega_n$ for all $0 \le i \le n$. In particular,

$$\left.\begin{aligned} W_i(\emptyset) &= 0 && (0 \le i \le n) \\ W_n(Y) &= \omega_n && (Y \in C/\{\emptyset\}) \\ W_0(Y) &= V(Y) && \\ W_1(Y) &= S(Y)/n && \end{aligned}\right\} . \tag{3.2}$$

By comparison of (3.1) with (2.11), by use of (2.15), and letting W_i^r denote the ith quermassintegral in E^r,

$$W_i(Y) = \omega_n \omega_{r-i} \int W_i^r\left(Y \cap F_r\right) dF_r / \omega_r \omega_{n-i} \int dL_r \quad (0 \le i \le r < n) . \tag{3.3}$$

Putting $r = n - 1$ (so that F_r is a hyperplane), we obtain

$$W_i(Y) = \omega_{n-i-1} \int W_i^{n-1}\left(Y \cap F_{n-1}\right) dF_{n-1} / n\omega_{n-1}\omega_{n-i} , \tag{3.4}$$

which is sometimes used to define the quermassintegrals recursively, together with the boundary conditions $W_n \equiv \omega_n$ and $W_0^1 = L$ (length).

The quermassintegrals possess a number of useful properties, for whose proof the reader is referred to Hadwiger (1957, Chapter 6).

(i) $W_i(Y)$ is invariant under translations and/or rotations of Y.

(ii) $W_i(aY) = |a|^{n-i} W_i(Y)$ for any scalar a, where $aY = \{a\mathrm{x} \mid \mathrm{x} \in Y\}$.

(iii) W_i is a continuous functional on $C/\{\emptyset\}$ with respect to the Hausdorff metric (i.e. the distance defined by $d\left(Y_1, Y_2\right) = \inf\{\varepsilon \ge 0 \mid Y_1 \subseteq Y_2 + \varepsilon O \text{ and } Y_2 \subseteq Y_1 + \varepsilon O\}$, where O is the unit n-ball).

(iv) W_i is additive, in the sense that if Y_1, Y_2 and $Y_1 \cup Y_2$ all belong to C, then

$$W_i(Y_1 \cup Y_2) + W_i(Y_1 \cap Y_2) = W_i(Y_1) + W_i(Y_2) . \tag{3.5}$$

In fact (i)-(iv) provide a characterization of W_i (Hadwiger (1957), p. 221).

Equation (3.5) suggests a way in which W_i could be extended to non-convex sets. Let K be the class of all finite unions of members of C. Any element Y of K may be written (although not uniquely) in the form $Y = \bigcup_{j=1}^{m} Y_j$, $Y_j \in C$. Consider the linear combination

$$\Psi = \sum_J (-1)^{|J|-1} W_i(Y_J) \quad (0 \le i \le n) \tag{3.6}$$

where J ranges over all non-empty subsets of $\{1, \ldots, m\}$, $|J|$ is the cardinality of J, and $Y_J = \bigcap_{j \in J} Y_j$. From (3.1),

$$\Psi = \left[\sum_J (-1)^{|J|-1} \omega_n \int_{F_i \uparrow Y_J} dF_i\right] / \omega_{n-i} \cdot \int dL_i$$

$$= \omega_n \int \left[\sum_J (-1)^{|J|-1} \chi(Y_J \cap F_i)\right] dF_i / \omega_{n-i} \int dL_i \quad (0 \le i \le n) , \tag{3.7}$$

where $\chi(A) = 1$ if $A \in C/\{\emptyset\}$ and $\chi(\emptyset) = 0$.

But the quantity $\sum_J (-1)^{|J|-1} \chi(Y_J \cap F_i)$ is independent of the particular representation chosen for Y (see Hadwiger (1957), p. 238), and may be written as $\chi(Y \cap F_i)$. χ is the Euler-Poincaré characteristic of differential geometry. Hence,

$$\Psi = \omega_n \int \chi(Y \cap F_i) dF_i / \omega_{n-i} \int dL_i , \tag{3.8}$$

which is also independent of the particular representation chosen for Y. We may therefore define $W_i(Y) = \Psi$, which is an additive extension of W_i from C to K. In fact (3.6) remains valid for $Y_j \in K$ instead of $Y_j \in C$.

The extended functional satisfies $W_0 \equiv V$, $W_1 \equiv S/n$ and $W_n \equiv \omega_n . \chi$. By linearity of integration, (3.3) remains valid for $Y \in K$. This equation plays a key role in stereology, as it relates r-dimensional information contained in cross-sections to n-dimensional properties of the entire set. The reader may wonder if K is a sufficiently broad class of sets within which to work. The answer is "yes" in the following sense. Matheron (1975, p. 117) has shown that K is dense (with respect to the Hausdorff metric) in the space of all compact sets. Hadwiger (1959) has extended the quermassintegrals to the so-called normal bodies, but we shall not consider this degree of generality. Other types of sets are treated in Sections 4-6.

One further useful equation, relating n- and r-dimensional quermassintegrals, is (Hadwiger (1957, p. 215))

$$W_i^r\left(Y \cap F_r\right) = \binom{n}{r-i}\omega_i W_{n+i-r}\left(Y \cap F_r\right) \Big/ \binom{r}{r-i}\omega_{n+i-r} \quad (0 \le i \le r < n) \ . \tag{3.9}$$

3.3. Centroids and vectorial integral geometry

Analogous to the real-valued functionals of the preceding section, we may define a family of vector-valued functionals $\mathsf{q}_i : \mathcal{C} \to E^n$, $0 \le i \le n$. Put $\mathsf{q}_i(\emptyset) = 0$ and $\mathsf{q}_n(Y) = \omega_n E[\lambda(\mathsf{u})\mathsf{u}]$, where u is an IR unit vector and $\lambda(\mathsf{u})$ is the value of the support function of Y in the direction u . For $0 \le i < n$, q_i is defined by

$$\mathsf{q}_i(Y) = \int \mathsf{q}_n\left(Y \cap F_i\right) dF_i / \omega_{n-i} \int dL_i^n \ . \tag{3.10}$$

Hadwiger and Schneider (1971) have shown that the functionals q_i satisfy a number of rather natural properties which we now state without proof.

(i)' q_i is equivariant under rotations and W_i-equivariant under translations. In other words,

$$\left.\begin{aligned} &\mathrm{q}_i[\tau(Y)] = \tau[\mathrm{q}_i(Y)] \quad \text{for all rotations } \tau \text{ of } E^n \\ \text{and}\quad & \\ &\mathrm{q}_i(Y+\mathrm{y}) = \mathrm{q}_i(Y) + W_i(Y).\mathrm{y} \quad (\mathrm{y} \in E^n) \end{aligned}\right\} . \tag{3.11}$$

(ii)' $\mathrm{q}_i(aY) = a|a|^{n-i}\mathrm{q}_i(Y) \quad (a \in \mathbb{R})$.

(iii)' q_i is continuous on $C/\{\emptyset\}$ with respect to the Hausdorff metric.

(iv)' If Y_1, Y_2 and $Y_1 \cup Y_2$ all belong to C , then

$$\mathrm{q}_i(Y_1 \cup Y_2) + \mathrm{q}(Y_1 \cap Y_2) = \mathrm{q}_i(Y_1) + \mathrm{q}_i(Y_2) . \tag{3.12}$$

Also,

$$\mathrm{q}_i(Y) = \binom{n}{r-i}\omega_n\omega_i\omega_{r-i} \int \mathrm{q}_{n+i-r}(Y \cap F_r)\,dF_r \Big/ \binom{r}{r-i}\omega_r\omega_{n+i-r}\omega_{n-i} \int dL_r$$

$$(0 \le i \le r < n) . \tag{3.13}$$

q_i may be extended additively to K via

$$\mathrm{q}_i(Y) = \sum_J (-1)^{|J|-1}\mathrm{q}_i(Y_J) \quad \left(Y = \bigcup_1^m Y_j,\ Y_j \in C\right) . \tag{3.14}$$

(As is the case for the quermassintegrals, the right-hand side is in fact independent of the particular representation chosen for Y .) (i)', (ii)', (iv)', (3.13) and (3.14) remain valid for all $Y, Y_j \in K$.

So far we have given no interpretation of q_i . Provided that $V(Y) \neq 0$, $\mathrm{q}_0(Y)/V(Y) = \int_Y \mathrm{x}d\mathrm{x}/\int_Y d\mathrm{x}$ is the centroid of Y . Similarly, we define

$$\mathrm{p}_i(Y) = \mathrm{q}_i(Y)/W_i(Y) \quad (0 \le i \le n,\ W_i(Y) \neq 0) . \tag{3.15}$$

For $i = 0$, p_i is the centroid of Y ; for $i = 1$, p_i is the centroid of the boundary ∂Y (i.e. $\mathrm{p}_1 = \int_{\partial Y} \mathrm{x}dS/S(Y)$, where dS is the element of surface measure at x); for $i = n$, p_i is the so-called Steiner point of Y (see Schreider (1971)). All of the p_i are measures of the location of Y . It may be shown, although it is not immediately obvious, that when $Y \in C/\{\emptyset\}$, $p_i \in Y$. For an n-ball, p_i is the centre of the ball for all i . If Y is contained within some r-flat, then $p_i^r(Y) = p_{n+i-r}(Y)$, where, as has been the convention up to now, the superscript denotes the dimension of the containing space and is usually omitted for dimension n $(0 \le i \le r < n)$.

Using (i)-(iv) and (i)'-(iv)', the following properties may be deduced for the functionals p_i . (It is supposed that $Y \in K$ and $W_i(Y) \neq 0$.)

(i)" p_i is equivariant under rotations and translations, i.e.

$$p_i[\tau(Y)] = \tau[p_i(Y)] \quad \text{for all rotations } \tau , \tag{3.16}$$

and

$$p_i[Y+y] = p_i(Y) + y \quad \text{for all } y \in E^n . \tag{3.17}$$

(ii)" $p_i(aY) = a p_i(Y) \quad (a \in \mathbb{R})$.

(iv)" p_i satisfies the *weighted* additivity property:

$$p_i(Y_1 \cup Y_2) =$$

$$= [W_i(Y_1)p_i(Y_1)+W_i(Y_2)p_i(Y_2)-W_i(Y_1 \cap Y_2)p(Y_1 \cap Y_2)]/W_i(Y_1 \cup Y_2) . \tag{3.18}$$

If we adopt the convention $W_i(\emptyset)p_i(\emptyset) = 0$, then (3.18) tells us that when Y consists of a number of disjoint particles, $p_i(Y)$ is equal to the W_i-weighted mean of the p_i's of the individual particles. When the particles are reduced to points, all of the p_i's vanish except for p_n , which is the centroid of the points.

3.4. Curvatures

We turn now to some apparently unrelated properties from differential geometry and show that these are in fact strongly related to the integral geometric functionals W_i, q_i and p_i .

The curvature κ of a smooth (twice continuously differentiable) curve in E^2 is defined by

$$dt = (\kappa ds)N \ , \tag{3.19}$$

where t, N are the unit tangent and normal vectors to the curve and s is arc length. (A convention must be adopted for the sense of N, but the sign of κ is independent of the sense chosen for t.)

When we consider a smooth $(n-1)$-dimensional hypersurface in E^{n-1}, the situation is somewhat more complicated. Suppose that the surface is oriented with unit normal vector N. If we intersect this surface with the (2-dimensional plane determined by N and a unit tangent vector u, we obtain a curve in E^2 which has an associated curvature $\kappa(u)$. $\kappa(u)$ varies as u rotates over all orientations within the tangent hyperplane. Classical differential geometry (see Hicks (1965)) shows that there exists a set of scalars $\{\kappa_1, \ldots, \kappa_{n-1}\}$ and an associated orthonormal basis $\{e_1, \ldots, e_{n-1}\}$ for the tangent hyperplane such that

$$\kappa(u) = \sum_{i=1}^{n-1} \kappa_i (e_i . u)^2 \ . \tag{3.20}$$

κ_i, e_i are in general unique to within a permutation of indices and sense reversal of e_i, but where some of the κ_i's are equal the associated e_i's may be chosen arbitrarily from the subspace which they span. The scalars κ_i are called the principal curvatures of the surface, and the e_i's are called the principal directions.

Consider the n symmetric functions of the principal curvatures

$$h_0 = 1$$

$$h_1 = \sum_{i=1}^{n-1} \kappa_i / n$$

$$\vdots \qquad (3.21)$$

$$h_i = \sum \kappa_{j_1} \cdots \kappa_{j_i} / \binom{n}{j}$$

(summation extending over all combinations of i indices from $\{1, \ldots, n-1\}$)

$$\vdots$$

$$h_{n-1} = \kappa_1 \ldots \kappa_{n-1} .$$

These quantities are invariant under permutations of the indices of the κ_i and are thus uniquely determined at each point on the surface.

For a bounded smooth surface F with element of surface measure dS, put

$$K_i(F) = \int_F h_i dS \quad (0 \le i < n) . \qquad (3.22)$$

K_0 is just the total surface area of F ; K_1 is the integral of mean curvature; K_n is the integral G of Gaussian curvature, which equals $\sigma_n = n\omega_n$ (the surface area of the unit n-sphere) when F is a simple closed surface.

It may be shown, using techniques similar to those employed by Miles (1975), that the following formula holds:

$$K_i(F) = \sigma_n \omega_{r-i-1} \int K_i^r(F \cap F_r) dF_r / \sigma_r \omega_{n-i-1} \int dL_r \quad (0 \le i < r < n) . \qquad (3.23)$$

Note the striking similarity between (3.23) and (3.3). In fact when F is the boundary ∂Y of a smooth compact convex set, the two equations reduce to the same thing. It has been shown by Bonnesen and Fenchel (1948) that for such a set Y ,

$$K_i(\partial Y) = nW_{i+1}(Y) \quad (0 \le i < n) \ . \tag{3.24}$$

For a general member of C, possibly possessing cusps, corners or edges, we may define K_i by

$$K_i(\partial Y) = \lim_{m\to\infty} K_i(\partial Y_m) \ , \tag{3.25}$$

where $\{Y_m\}$ is a sequence of sets having twice continuously differentiable boundaries such that $Y = \lim_{m\to\infty} Y_m$ (see Matheron (1975, p. 114)). For example, if Y is a convex polytope with m $(n-i-1)$-dimensional facets Y_j having exterior angles ψ_j, then

$$K_i(\partial Y) = \sum_{j=1}^{m} V_{n-i-1}(Y_j)\psi_j \ . \tag{3.26}$$

In effect, we are replacing the symmetric function of principal curvatures by a generalized function over the boundary of Y (which is concentrated on the $(n-i-1)$-facets in the case where Y is a polytope). Matheron (1975, p. 113) has rigorized this notion. Using h_j to denote the generalized function, we may define

$$K_i(\partial Y) = \int_{\partial Y} \left[\sum_{J} (-1)^{|J|-1} h_i(\partial Y_J)\right] dS \tag{3.27}$$

for $Y = \bigcup_1^m Y_j$, $Y_j \in C$. To see that this is a proper definition (i.e. that it does not depend upon the representation chosen for Y), first observe that by (3.25) and the continuity of W_{i+1}, (3.24) is valid for all $Y \in C$. Therefore

$$\int_{\partial Y}\left[\sum_{J}(-1)^{|J|-1}h_i(\partial Y_J)\right]dS = \sum_{J}(-1)^{|J|-1}\int_{\partial Y_J} h_i(\partial Y_J)dS$$

$$= \sum_{J}(-1)^{|J|-1}K_i(\partial Y_J)$$

$$= n\sum_{J}(-1)^{|J|-1}W_{i+1}(Y_J)$$

$$= n.W_{i+1}(Y) \ . \tag{3.28}$$

Thus the quermassintegrals W_i may be interpreted as integrals of generalized curvature for $i > 0$.

The functionals p_{i+1} have a similar representation

$$\mathrm{p}_{i+1}(Y) = \int_{\partial Y} h_i x dS \Big/ \int_{\partial Y} h_i dS \quad (0 \le i < n,\ Y \in C/\{\emptyset\}) \ . \tag{3.29}$$

3.5. Mean projections

It has already been observed that the total measure of r-flats hitting $Y \in C$ may be expressed both in terms of the mean projection $M_{n-r}(Y)$ of Y onto an IR $(n-r)$-subspace and in terms of the rth quermassintegral $W_r(Y)$. By comparison of (3.1) and (2.11),

$$W_r(Y) = \omega_n M_{n-r}(Y)/\omega_{n-r} \quad (0 \le r \le n) \ . \tag{3.30}$$

This identity fails to hold for all $Y \in K$, but can be replaced by

$$W_r(Y) = \omega_n \int \left(\int_{Y|L_{n-r}} \chi(Y \cap F_r)dx_{n-r}\right)dL_{n-r}/\omega_{n-r}\int dL_{n-r}$$

(where F_r is the r-flat orthogonal to L_{n-r} and passing through x_{n-r})

$$= \omega_n \sum_{J}(-1)^{|J|-1}\int\int_{Y_J|L_{n-r}} dx_{n-r}dL_{n-r}/\omega_{n-r}\int dL_{n-r}$$

(where $Y = \bigcup_1^m Y_j$, $Y_j \in C$)

$$= \sum_{J}(-1)^{|J|-1}\omega_n M_{n-r}(Y_J)/\omega_{n-r} \ . \tag{3.31}$$

Hence W_r may be interpreted both as a mean χ-weighted projection and as a linear combination of mean projections of convex subsets.

3.6. Fractals and varieties

So far we have considered certain types of $(n-1)$-dimensional surfaces, but not varieties of dimension lower than $n - 2$, such as curves in E^3. Nor have we considered fractals, i.e. sets of fractional dimension, which have been used recently by Mandelbrot (1977) to describe phenomena such as coastlines, lunar surfaces and porous materials.

There is a variety of measures which have been defined for varieties and fractals. Most of these are based on Carathéodory's construction (see Federer (1969, p. 169)), which can be described loosely as follows. A set is covered as economically as possible from a collection of "tiles" of bounded maximum diameter, the sizes of the tiles used are added up, and then the supremum is taken as the diameter bound for the available tiles tends to zero. Different measures are obtained by varying the collection of tiles and/or the definition of their sizes. These Carathéodory measures apply to any Borel set (in particular there are no smoothness or convexity requirements), but may attain the value ∞. The best known of them is m-dimensional Hausdorff measure H^m, for which the tiles consist of all non-empty subsets of E^n and size is proportional to the mth power of maximum diameter. The dimension of a set Y is the infimum of m such that $H^m(Y) > 0$.

Unfortunately, certain pathological examples cause difficulties in the application of stereology to the above measures. For example, it is not always true that the general intersection of an r-flat with a set of Hausdorff dimension m yields a set of dimension $m + r - n$. However there is one form which lends itself well to integral geometric formulae.

This is the m-dimensional integralgeometric measure of exponent 1 , defined by

$$I^m(Y) = \sigma_{m+1}\sigma_{n-m+1} \sup_{\delta>0} \inf_{G(\delta)} \sum_{S\in G(\delta)} M_m(S)/2\sigma_{n+1} \quad (0 \le m \le n) \quad (3.32)$$

where Y is a Borel set and $G(\delta)$ ranges over all countable families of compact sets S whose maximum diameter is less than δ .

Federer has shown that the following formula holds for r-flats hitting Y :

$$I^m(Y) = \sigma_{m+1}\sigma_{n-m+1} \int I^{m+r-n}\left(Y \cap F_r\right) dF_r/2\sigma_{n+1} \int dL_r \quad (n-r \le m \le n) \ . \quad (3.33)$$

Furthermore, for the class of $\left(H^m, m\right)$ rectifiable subsets of E^n (Federer (1969, p. 261)), I^m and H^m coincide.

Let us look at some particular examples. I^n is just n-dimensional Lebesgue measure. When Y is a unit m-ball regarded as a subset of E^n , $I^m(Y) = \omega_m$. For the boundary of a set $Y \in K$,

$$I^{n-1}(\partial Y) = K_0(\partial Y) = S(Y) = nW_1(Y) \ . \quad (3.34)$$

Finally, when Y can be expressed as a finite union of m-dimensional compact convex sets, then

$$I^m(Y) = \sigma_{m+1}\sigma_{n-m+1}\omega_m W_{n-m}(Y)/2\sigma_{n+1}\omega_n \ . \quad (3.35)$$

From the above comments, it can be seen that the equations (3.3), (3.23) and (3.33) are equivalent in the following cases:

(3.3) $(i = 0)$ is equivalent to (3.33) $(m = n)$.

(3.3) $(i = 1)$ is equivalent to (3.23) $(i = 0,\ F = \partial Y)$ and to (3.33) $(m = n-1,\ Y = \partial Y)$.

(3.3) $(i = m)$ is equivalent to (3.33) (Y a finite union of m-dimensional compact convex sets).

3.7. Functions and vector fields

Suppose that there is an unknown, measurable real-valued function $f : X \to \mathbb{R}$ or vector-valued function $\mathrm{f} : X \to E^n$ and that we are interested in certain global characteristics of this function.

By using (2.9) with $q = 0$,

$$\int \left(\int_{X \cap F_r} f d\mathrm{x}_r \right) dF_r = \int_X \left(\int f dF_{r(0)} \right) d\mathrm{x} = \int dL_r \cdot \int_X f d\mathrm{x} \quad (0 \le r < n) . \tag{3.36}$$

The term enclosed in brackets on the left-hand side of (3.36), regarded as a function of the flat F_r, is known as the Radon transform of f (see Gel'fand, Graev and Vilenkin (1966)). The equation states that integration of f over X may be performed by the integration of its Radon transform over all flats hitting X.

We may derive a similar expression involving f_r, the projection of f onto F_r (i.e. $\mathrm{f}_r(\mathrm{x})$ is equal to the projection of the vector $\mathrm{f}(\mathrm{x})$ onto F_r for all $\mathrm{x} \in F_r$) as follows. First,

$$\int \left(\int_{X \cap F_r} \mathrm{f}_r d\mathrm{x}_r \right) dF_r = \int_X \left(\int \mathrm{f}_r dF_{r(0)} \right) d\mathrm{x} . \tag{3.37}$$

To evaluate the term within brackets on the right-hand side of (3.37), recall that a vector $\mathrm{y} \in E^n$ may be written as $\mathrm{y} = \sum_1^n (\mathrm{y}.\mathrm{e}_i)\mathrm{e}_i$ for an arbitrary orthonormal basis $\{\mathrm{e}_i\}$. If we subject $\{\mathrm{e}_i\}$ to an IR rotation, and take expectations, we find that

$$\mathrm{y} = nE[(\mathrm{y}.\mathrm{e})\mathrm{e}] , \tag{3.38}$$

where e is an IR unit vector in E^n. Similarly, the projected vector $\mathrm{y}_r = rE[(\mathrm{y}.\mathrm{e}_r)\mathrm{e}_r]$ where e_r is an IR unit vector parallel to F_r. Hence

$$\int y_r \, dF_{r(0)} = r \int E[(y.e_r)e_r] \, dF_{r(0)}$$

$$= r \int dL_r E[(y.e)e] \quad \text{by isotropy}$$

$$= \frac{r}{n} \int dL_r y \ . \tag{3.39}$$

Applying this to (3.37) yields

$$\int \left(\int_{X \cap F_r} f_r \, dx_r \right) dF_r = \frac{r}{n} \int dL_r \int_X f dx \ . \tag{3.40}$$

Replacing f_r by its modulus,

$$\int \left(\int_{X \cap F_r} |f_r| \, dx_r \right) dF_r = \int_X \left(\int |f_r| \, dF_{r(0)} \right) dx$$

$$= \lambda \int dL_r \ . \int_X |f| dx \ , \tag{3.41}$$

where λ is the mean length of projection of a fixed unit vector onto an isotropic random r-subspace. By symmetry, λ is equal to half the mean projection $M_1(O_r) = 2W_{n-1}(O_r)/\omega_n$. From Hadwiger (1957), $W_{n-1}(O_r) = \sigma_r \omega_{n-1}/n\omega_{r-1}$, and hence $\lambda = \sigma_r \omega_{n-1}/\sigma_n \omega_{r-1}$.

If f, f are differentiable, we may consider the operators grad, div and rot, denoted by ∇f, $\nabla . \mathrm{f}$ and $\nabla \times \mathrm{f}$ respectively. By the invariance of these operators with respect to the (rectangular) co-ordinate system used, we may proceed as above and write

$$\nabla f = nE\left[\frac{\partial f}{\partial s} \mathrm{e}\right] \qquad \nabla_r f = rE\left[\frac{\partial f}{\partial s_r} \mathrm{e}_r\right]$$

$$\nabla . \mathrm{f} = nE\left[\frac{\partial}{\partial s} (\mathrm{f}.\mathrm{e})\right] \qquad \nabla_r . \mathrm{f}_r = rE\left[\frac{\partial}{\partial s_r} (\mathrm{f}.\mathrm{e}_r)\right]$$

$$\nabla \times \mathrm{f} = nE\left[\frac{\partial}{\partial s} (\mathrm{f} \times \mathrm{e})\right] \qquad \nabla_r \times \mathrm{f}_r = rE\left[\frac{\partial}{\partial s} (\mathrm{f}_r \times \mathrm{e}_r)\right]$$

$$= n(n-1)E\left[\frac{\partial}{\partial s} (\mathrm{f}.\mathrm{d})\mathrm{d} \times \mathrm{e}\right] \qquad = r(r-1)E\left[\frac{\partial}{\partial s_r} (\mathrm{f}.\mathrm{d}_r)\mathrm{d}_r \times \mathrm{e}_r\right] \ . \tag{3.42}$$

Here the subscript r on operators signifies that they are applied within F_r , e (e_r) is an IR unit vector in E^n (F_r) , s (s_r) is arc length along e (e_r) and d (d_r) is an IR unit vector orthogonal to e (e_r) .

From (3.41), it follows easily that

$$\int \left(\int_{X \cap F_r} \nabla_r f dx_r \right) dF_r = \frac{r}{n} \int dL_r \int_X \nabla f dx \ , \tag{3.43}$$

$$\int \left(\int_{X \cap F_r} \nabla_r . f_r dx_r \right) dF_r = \frac{r}{n} \int dL_r \int_X \nabla . f dx \ , \tag{3.44}$$

$$\int \left(\int_{X \cap F_r} \nabla_r \times f_r dx_r \right) dF_r = \frac{r(r-1)}{n(n-1)} \int dL_r \int \nabla \times f dx \ . \tag{3.45}$$

In the case when X is bounded by a smooth hypersurface, we may use classical vector analysis to deduce from (3.43) and (3.44)

$$\int \left(\int_{\partial X \cap F_r} N_r . f_r dS_r \right) dF_r = \frac{r}{n} \int dL_r \int_{\partial X} N . f dS \tag{3.46}$$

and

$$\int \left(\int_{\partial X \cap F_r} N_r \times f_r dS_r \right) dF_r = \frac{r(r-1)}{n(n-1)} \int dL_r \int_{\partial X} N \times f dS \ , \tag{3.47}$$

where N (N_r) is the unit outward normal to ∂X $(\partial X \cap F_r)$.

By observing that $\nabla_r f$ is actually the projection of ∇f onto F_r , we may obtain from (3.40),

$$\int \left(\int_{X \cap F_r} |\nabla_r f| dx_r \right) dF_r = \frac{\sigma_r \omega_{n-1}}{\sigma_n \omega_{r-1}} \int dL_r \int_X |\nabla f| dx \ . \tag{3.48}$$

The above technique can be applied to higher order operators: for example

$$\int \left(\int_{X \cap F_r} \nabla_r^2 f dx_r \right) dF_r = \frac{r}{n} \int dL_r \int_X \nabla^2 f dx \tag{3.49}$$

which in the case when X is enclosed by a smooth boundary can be written as

$$\int \left(\int_{\partial X \cap F_r} \frac{\partial f}{\partial N_r} dS_r \right) dF_r = \frac{r}{n} \int dL_r \int_{\partial X} \frac{\partial f}{\partial N} dS \ , \tag{3.50}$$

the partial derivatives being evaluated along the outward normals.

CHAPTER 4

ESTIMATORS AND THEIR VARIANCES

4.1. Introduction

In this chapter we go systematically through the geometrical properties described in Chapter 3 and present various unbiased estimators associated with the sampling schemes of Chapter 2. Variances are also derived, although the expressions for some of these are a little unwieldy due to the generality allowed for the structure of the speimen. Where a number of estimators exist for the same property, some comparisons of efficiency have been made. Each section is sprinkled liberally with examples in two and three dimensions.

4.2. Quermassintegrals

Let us suppose that the feature set $Y \subset X$ is a member of K. From (2.11) and (3.3), it follows that for an IUR r-flat F_r hitting X,

$$E\left[\omega_n \omega_{r-i} M_{n-r}(X) W_i^r\left(Y \cap F_r\right) / \omega_r \omega_{n-i}\right] = W_i(Y) \quad (0 \leq i \leq r < n) . \tag{4.1}$$

The quantity α_r within square brackets on the left-hand side of (4.1) depends upon the external property $M_{n-r}(X)$ which is supposed known and the cross-sectional property $W_i^r\left(Y \cap F_r\right)$ which is observable. α_r is an unbiased estimator of $W_i(Y)$ (assumed to be unknown).

To illustrate (4.1), take $n = 3$, $r = 2$, $i = 0, 1, 2$. We obtain

$$E\left[M(X) A\left(Y \cap F_2\right)\right] = V(Y) , \tag{4.2}$$

$$E\left[4M(X) B\left(Y \cap F_2\right) / \pi\right] = S(Y) , \tag{4.3}$$

$$E\left[2\pi M(X) \chi\left(Y \cap F_2\right)\right] = K(Y) , \tag{4.4}$$

where A, B denote area and boundary perimeter and M, K denote mean linear projection (M_1^3) and integral of mean curvature (K_1^3).

In order to evaluate the variance of α_r, we need the following integral geometric identities.

THEOREM 4.1. *For* $2i < r < n$,

$$dF_i^r dF_i'^r dF_r = \delta^{r-n} \sin^{r-n}\theta dF_{r(2i+1)} dF_i dF_i' . \tag{4.5}$$

A little explanation of notation is needed here. On the left-hand side we have the joint element of two i-flats contained within an r-flat. On the right-hand side, $F_{r(2i+1)}$ is the r-flat containing the two i-flats F_i and F_i', which span a $(2i+1)$-flat. δ is the distance between F_i and F_i', and θ is the angle between the subspaces L_i, L_i' parallel to F_i and F_i', defined as follows. Let L_{2i} be the subspace spanned by L_i and L_i', and L_i'' the orthogonal complement of L_i in L_{2i}. $\sin\theta$ is equal to the i-volume of the projection onto L_i'' of a set Z of unit i-volume in L_i'. Miles (1972a) has demonstrated that $\sin\theta$ is independent of the particular set Z chosen and that the definition is symmetric with respect to L_i and L_i'. Note that for $1 < i < r-1$ this is not in accordance with the usual treatment of angles between subspaces (see Kendall (1961), for example). However in the familiar cases of 2 and 3 dimensions, θ corresponds to the usual angle between lines. For convenience, we take $\sin\theta$ to be 1 when $i = 0$.

Proof. From results of Miles (1971a), it may be verified that

$$dF_i dF_i' = \delta^{n-2i-1} \sin^{n-2i-1}\theta dF_i^{2i+1} dF_i'^{2i+1} dF_{2i+1} . \tag{4.6}$$

Hence

$$dF_{r(2i+1)}dF_i.dF_i' = \delta^{n-2i-1}\sin^{n-2i-1}\theta dF_i^{2i+1}dF_i'^{2i+1}dF_{2i+1}^r dF_r \quad \text{(from (2.9))}$$

$$= \delta^{n-r}\sin^{n-r}\theta dF_i^r.dF_i'^r dF_r \quad \text{(from (4.6), putting } n = r\text{)}. \quad //$$

THEOREM 4.2. *For* $2i \geq r$, $r < n$,

$$dF_i^r.dF_i'^r dF_r = \sin^{2i+1-n}\theta dF_{i(2i-r)}dF'_{i(2i-r)}dF_{2i-r} . \tag{4.7}$$

Here $F_{i(2i-r)}$, $F'_{i(2i-r)}$ are two IR i-flats containing their intersection F_{2i-r}. θ is the angle between the subspaces L_{r-i}, L'_{r-i} parallel to F_r and orthogonal to F_i, F_i'.

Proof. From equation (12.47) of Santaló (1976),

$$dF_i^r.dF_i'^r = \sin^{2i+1-r}\theta dF_{i(2i-r)}^r dF'^r_{i(2i-r)}dF_{2i-r}^r .$$

Hence

$$dF_i^r.dF_i'^r dF_r = \sin^{2i+1-r}\theta dF_{i(2i-r)}^r dF'^r_{i(2i-r)}dF_{r(2i-r)}dF_{2i-r} \quad \text{(from (2.9))}$$

$$= \sin^{2i+1-r}\theta\left(dL_{r-i}^{2r-2i}\,dL'^{2r-2i}_{r-i}\,dL_{2r-2i}^{n-2i+r}\right)dF_{2i-r} \quad \text{(from (2.3))}.$$

But from Miles (1971a), the term within brackets is equal to

$\sin^{r-n}\theta dL_{r-i}^{n-2i+r}\,dL'^{n-2i+r}_{r-i}$, and hence

$$dF_i^r.dF_i'^r dF_r = \sin^{2i+1-n}\theta dF_{i(2i-r)}^n dF'^n_{i(2i-r)}dF_{2i-r} \quad \text{(from (2.3))}. \quad //$$

Multiplying both sides of (4.5) and (4.7) by $\chi\chi' = \chi(Y \cap F_i)\chi(Y \cap F_i')$ and integrating over the set of configurations such that F_i and F_i' both hit X, we obtain after slight rearrangement

$$E\left[(\alpha_r)^2\right] = \left(\frac{\omega_n}{\omega_{n-i}}\right)^2 \frac{M_{n-r}(X)}{\int dL_r\left(\int dL_i^r\right)^2}$$

$$\times\begin{cases}\displaystyle\iint dL_{r(2i+1)}\iint \chi\chi'\delta^{r-n}\sin^{r-n}\theta dF_i.dF_i' & (2i < r) \\ \displaystyle\iiint \chi\chi'\sin^{2i+1-n}\theta dF_{i(2i-r)}dF'_{i(2i-r)}dF_{2i-r} & (2i \geq r) .\end{cases} \tag{4.8}$$

By subtracting $[W_i(Y)]^2$ from both sides of (4.8) we obtain the variance of the estimator α_r .

As a first example, put $i = 0$. Then α_r is an estimator of the volume of Y , and

$$\mathrm{Var}(\alpha_r) = \begin{cases} [V(Y)]^2\left\{\frac{\sigma_r}{\sigma_n} M_{n-r}(X).E\left(\delta^{r-n}\right)-1\right\} & (r > 0) \\ \\ [V(Y)]^2\{V(X)/V(Y)-1\} & (r = 0) \end{cases} , \tag{4.9}$$

where $E\left(\delta^{r-n}\right)$ is the $(r-n)$th moment of the distance between two independent UR points of Y . An intuitive interpretation of (4.9) is that the variance of α_r is small when the specimen X is small and Y is well "spread out" through X .

As a second example, put $n = 3$, $r = 2$, $i = 1$. Then α_2 is an unbiased estimator of $S(Y)/3$, and

$$\mathrm{Var}(\alpha_2) = \frac{32}{9\pi} M(X)V(Y) + \frac{8}{9\pi^3} M(X) \int_{Y^C} \omega^2 dx - \tfrac{1}{9}S^2(Y) , \tag{4.10}$$

where ω is the "total solid angle" subtended by Y at x (i.e. the total measure of lines F_1 passing through x weighted according to the number of disjoint intervals in $F_1 \cap Y$).

We have presented $(n-i)$ estimators $\alpha_i, \alpha_{i+1}, \ldots, \alpha_{n-1}$ for the property $W_i(Y)$. In comparing them, we should consider both ease of implementation and size of variance. For example, for $n = 3$, $i = 0$, $r = 0, 1, 2$, we are estimating $V(Y)$ via presence or absence of Y on a UR point, length of intersection of Y with an IUR line, and area of intersection of Y with an IUR plane respectively. The optimal sample dimension r on the criterion of variance alone is that which minimizes $\sigma_r M_{n-r}(X)E\left(\delta^{r-n}\right)/\sigma_n$ $(r > 0)$ or $V(X) - V(Y)$ $(r = 0)$. In the case when

Y and X are spheres of radii s and t $(s \leq t)$ respectively, $E(\delta^{-1}) = 6/5s$ and $E(\delta^{-2}) = 9/4s^2$. Hence, from (4.9) the relative variance of α_r is

r	$\mathrm{Var}(\alpha_r)/[V(Y)]^2$
0	$t^3/s^3 - 1$
1	$9t^2/8s^2 - 1$
2	$6t/5s - 1$.

This relative variance is plotted vs . t/s in Figure 4.1.

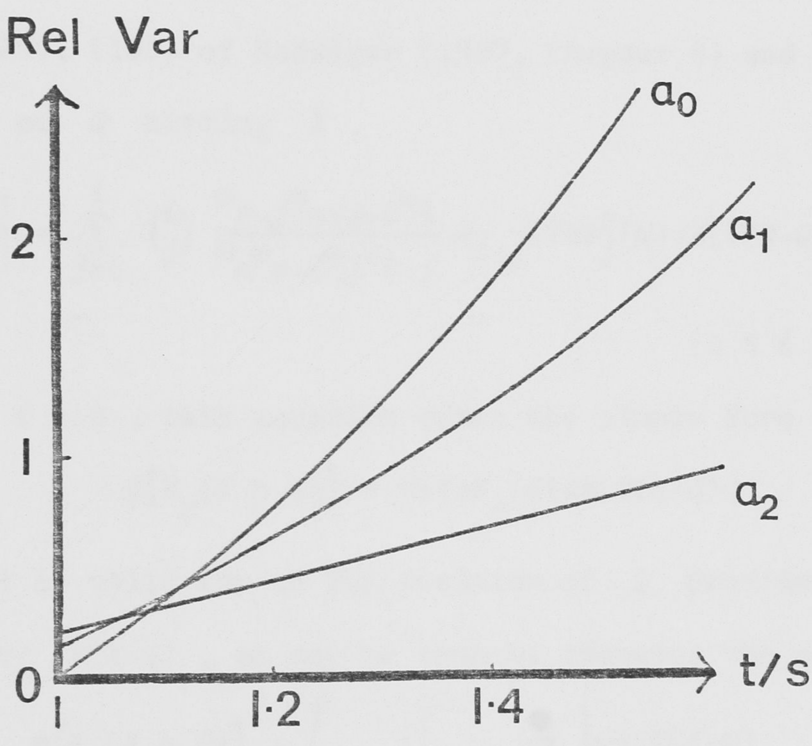


FIGURE 4.1. Relative variance of volume estimators for sphere of radius s embedded in sphere of radius t .

Note that the relative efficiencies of the three estimators change with increasing t/s . At first, the order of increasing variance for the sample dimension r is $(0, 1, 2)$. This ordering then changes to $(0, 2, 1)$, to $(2, 0, 1)$, and finally to $(2, 1, 0)$.

We have seen that estimation of the $r + 1$ quantities $W_0(Y), \ldots, W_r(Y)$ is possible on the basis of an r-section. Is it possible to estimate $W_{r+1}(Y), \ldots, W_n(Y)$ also? The answer in general is no. A unit $(n-i)$-ball O_{n-i} in E^n has non-zero W_i, and yet the measure of r-flats hitting O_{n-i} is zero for $r < i$. Hence, in the case $Y = O_{n-i}$, an IUR r-flat through X contains no information concerning $W_i(Y)$ for $i > r$.

When Y is convex, $W_i(Y)$ is proportional to $M_{n-i}(Y)$, and may be estimated by projectional as well as sectional methods (see Section 5).

The third type of sampling described in Chapter 2 involved quadrats. If the quadrat $Q \in K$ is r-dimensional, then it may be shown using equations (129), (130) of Hadwiger (1957, Chapter 6) and (2.28) that for an IUR position of Q hitting X,

$$E\left[W_i^r(Y \cap Q)\right] = \sum_{j=0}^{i} \binom{i}{j} \frac{\omega_{r-j}\omega_{n+j-i}\omega_i}{\omega_n\omega_{r-i}\omega_j\omega_{i-j}} W_{i-j}(Y)W_j^r(Q)/E[V(X-Q)] \qquad (0 \leq i \leq r \leq n) . \quad (4.11)$$

In the case $i = 0$, this equation takes the simple form

$$E\left[V_r(Y \cap Q)\right] = V(Y)V_r(Q)/E[V(X-Q)] . \quad (4.12)$$

In fact (4.8) is valid for an FUR position of Q provided that $E[V(X-Q)]$ is replaced by $V(X-Q)$, as may be seen by changing the order of integration.

$$\begin{aligned} E\left[V_r(Y \cap Q)\right] &= \int_{X-Q} \left(\int_{Y \cap Q_y} dx_r\right) dy / V(X-Q) \\ &= \int_{Y} \left(\int_{Q_y \ni x} dy_r\right) dx_r dy_{n-r} / V(X-Q) \\ &= V(Y)V_r(Q)/V(X-Q) . \end{aligned} \quad (4.13)$$

For $i > 0$, (4.11) cannot be used directly for the estimation of the quermassintegrals of Y, as a linear combination of the latter, rather than a single term, appears on the right-hand side. It turns out that the terms for $j > 0$ correspond to edge effects during sampling. This problem may be overcome in the following fashion. In Section 3.4, we observed that the quermassintegrals could be interpreted as surface integrals of generalized curvature for $i > 0$, i.e. $W_i(Y) = \int h_{i-1}(\partial Y)dS/n$. Instead of measuring $W_i^r(Y \cap Q)$ for a random position of Q, we may measure $\int h_{i-1}^r(\partial Y \cap Q)dS_r$. For example, taking $r = 2$ and $i = 1$, $W_1^2(Y \cap Q)$ is half the perimeter of $Y \cap Q$, while $\int h_0^2(\partial Y \cap Q)dS_2$ is half the length of $Q \cap \partial Y$.

When $r = n$, we may use a change in order of integration to prove that for an FUR position of Q hitting X,

$$E\left[\int h_{i-1}(\partial Y \cap Q)dS\right] = W_i(Y)V(Q)/V(X-Q) \quad (0 < i \leq n) . \tag{4.14}$$

By randomizing orientation, we find that (4.14) remains valid for an IUR quadrat of non-zero n-volume. In the case $r < n$, we can use the fact that an IUR quadrat has element $dQ^r dF_r / \int dB_r E[V(X-Q)]$, where dQ^r is the element of r-dimensional kinematic measure (see Santaló (1976, p. 258)). Hence (4.1) and (4.14) may be combined to obtain

$$E\left[\int h_{i-1}^r(\partial Y \cap Q)dS_r\right] = \frac{\omega_r \omega_{n-i}}{\omega_n \omega_{r-i}} W_i(Y)V_r(Q)/E[V(X-Q)] \quad (0 < i \leq r) . \tag{4.15}$$

Only the term $j = 0$ is retained from the right-hand side of (4.11).

Expressions for the variances of quadrat estimators may be derived in a similar way to that used for sectional estimators. We give one such example here.

PROPOSITION 4.3. *The variance of the unbiased estimator* $\alpha = V_r(Y \cap Q)E[V(X-Q)]/V_r(Q)$ *of* $V(Y)$ *is given by*

$$\mathrm{Var}(\alpha) = [V(Y)]^2\left\{E[V(X-Q)] \int_0^\infty \delta^{1-n} p_Y(\delta)p_Q(\delta)d\delta/\sigma_n - 1\right\} \tag{4.16}$$

where $p_Y(\cdot)$ (*resp.* $p_Q(\cdot)$) *is the probability density associated with the distance between two independent* UR *points of* Y (*resp.* Q).

Proof.

$$E[V_r(Y \cap Q)]^2 = \int_{Q\uparrow X}\int_{Y\cap Q}\int_{Y\cap Q} dx_r dx'_r \Big/ \int dB_r . E[V(X-Q)]$$

$$= \int_Y \int_Y \delta^{r-n}\left(\int_{Q\ni 0,\delta x} dQ\right) dx dx' \Big/ \int dB_r E[V(X-Q)]$$

$$= [V_r(Q)]^2 \int_Y \int_Y \delta^{r-n}\delta^{1-r}\left(\frac{\int dB_{r(1)}}{\sigma_r \int dB_r}\right) p_Q(\delta) dx dx'/E[V(X-Q)]$$

$$= [V_r(Y)]^2 [V_r(Q)]^2 \int_0^\infty \delta^{1-n} p_Q(\delta) p_Y(\delta) d\delta/\sigma_n E[V(X-Q)] \; .$$

The proposition follows upon multiplying both sides by $\{E[V(X-Q)]/V_r(Q)\}^2$ and subtracting $(E\alpha)^2$. //

4.3. Centroids

Suppose, as in the last section, that $Y \in K$. From (3.13), it follows that

$$\alpha_r = \frac{\binom{n}{r-i}\omega_n\omega_i\omega_{r-i}}{\binom{r}{r-i}\omega_r\omega_{n+i-r}\omega_{n-i}} M_{n-r}(X) q_{n+i-r}(Y \cap F_r) \quad (0 \le i \le r < n) \quad (4.17)$$

is an unbiased estimator of $q_i(Y)$ based on an IUR cross-section $Y \cap F_r$. There is an important difference between this estimator and those of Section 4.2 - here the value depends upon the *position* of F_r as well as upon information contained wholly within F_r . By use of (4.5) and (4.7) it may be shown that

$$E|\alpha_r - q_i(Y)|^2 = -|q_i(Y)|^2 + \frac{M_{n-r}(X)}{\omega_{n-i}^2 \int dL_r \left(\int dL_i^r\right)^2}$$

$$\times \begin{cases} \iiint q_n . q'_n \delta^{r-n} \sin^{r-n}\theta dF_i dF'_i \int dL_{r(2i+1)} & (2i < r) \\ \iiint q_n . q'_n \sin^{2i+1-n}\theta dF_{i(2i-r)} dF'_{i(2i-r)} dF_{2i-r} & (2i \ge r) \; , \end{cases} \quad (4.18)$$

where $q_n = q_n(Y \cap F_i)$, $q'_n = q_n(Y \cap F'_i)$.

If $p_i(Y)$ is of interest rather than $q_i(Y)$, we must divide α_r by $W_i(Y)$ to obtain the appropriate estimator. Generally $W_i(Y)$ will also be unknown. If it is replaced by its estimator, we obtain the (biased in general) estimator $q_{n+i-r}(Y \cap F_r)/W_i^r(Y \cap F_r)$ for $p_i(Y)$. In the next chapter it will be shown how this bias can be eliminated.

Suppose that the external characteristics $V(X)$, $S(X)$ and $M(X)$ of the 3-dimensional convex set X are known. Then from (4.17), the centroid c can be estimated either by generating a UR point x of X (estimator $\alpha_0 = x$), by generating an IUR secant through X (estimator $\alpha_1 = S(X) \int_{F_1 \cap X} x dx/4V(X)$) or by generating an IUR planar section through X (estimator $\alpha_1 = M(X) \int_{F_2 \cap X} x dx_2/V(X)$). The mean square errors of these three estimators are

$$\mathrm{MSE}(\alpha_0) = \frac{1}{V(X)} \int_X |x|^2 dx - |c|^2 , \tag{4.19}$$

$$\mathrm{MSE}(\alpha_1) = \frac{S(X)}{8\pi[V(X)]^2} \int_X \int_X \frac{x.x'}{|x-x'|^2} dx dx' - |c|^2 , \tag{4.20}$$

$$\mathrm{MSE}(\alpha_2) = \frac{M(X)}{2\pi[V(X)]^2} \int_X \int_X \frac{x.x'}{|x-x'|} dx dx' - |c|^2 . \tag{4.21}$$

4.4. Curvatures

If ∂Y is a smooth surface embedded in X , then from (3.23),

$$\alpha_r = \frac{\sigma_n \omega_{r-j-1}}{\sigma_r \omega_{n-j-1}} M_{n-r}(X) K_j^r(\partial Y \cap F_r) \quad (0 \le j < r < n) \tag{4.22}$$

is an unbiased estimator of $K_j(\partial Y)$. By (4.5) and (4.7), its variance is given by

$$\mathrm{Var}\left(\alpha_r\right) = -\left[K_j(\partial Y)\right]^2 + \left(\frac{\sigma_n}{\sigma_{j+1}\omega_{n-j-1}}\right)^2 \frac{M_{n-r}(X)}{\int dL_r\left(\int dL^r_{j+1}\right)^2}$$

$$\times \begin{cases} \iint dL_{r(2j+3)} \int KK'\delta^{r-n} \sin^{r-n}\theta dF_{j+1} dF'_{j+1} & (2j < r-2) \\ \iiint KK' \sin^{2j+3-n}\theta dF_{j+1(2j+2-r)} dF_{j+1(2j+2-r)} dF_{2j+2-r} & (2j \geq r-2) \end{cases} , \tag{4.23}$$

where $K = K_j^{j+1}\left(\partial Y \cap F_{j+1}\right)$, $K' = K_j^{j+1}\left(\partial Y \cap F'_{j+1}\right)$.

Putting $n = 3$, $r = 2$, $j = 1$ gives that $M(X)K_1^2\left(\partial Y \cap F_2\right)$ is an unbiased estimator of $K_1^3(\partial Y)$. $K_1^2\left(\partial Y \cap F_2\right)$ is the total curvature C of the curve $\partial Y \cap F_2$, while $K_1^3(\partial Y)$ is the integral of mean curvature K over the surface of Y .

Curvatures may also be estimated via quadrats. For example, if Q is r-dimensional, it may be shown in a way similar to the derivation of (4.15) that

$$\frac{\sigma_n\omega_{r-j-1}}{\sigma_r\omega_{n-j-1}} \frac{E[V(X-Q)]}{V_r(Q)} K_j^r(\partial Y \cap Q) \tag{4.24}$$

is an unbiased estimator of $K_j(\partial Y)$ for an IUR position of Q hitting X .

4.5. Mean projections

Let Y be a closed subset of X . By (2.15), (4.5) and (4.7),

$$M_{n-r}(X)M^r_{r+i-n}\left(Y \cap F_r\right) \quad (n-r \leq i \leq n) \tag{4.25}$$

is an unbiased estimator of $M_i(Y)$ for an IUR r-section of X , having variance

$$-\left[M_i(Y)\right]^2 + \left(\frac{\omega_i}{\omega_{r+i-n}}\right)^2 \frac{M_{n-r}(X)}{\int dL_r \left(\int dL_i^r\right)^2}$$

$$\times \begin{cases} \displaystyle\iint dL_{r(2n-2i+1)} \iint_{F_{n-i}, F'_{n-i} \uparrow Y} \delta^{r-n} \sin^{r-n}\theta dF_{n-i} dF'_{n-i} & (2i > 2n-r) \\ \displaystyle\iiint_{F_{n-i}, F'_{n-i} \uparrow Y} \sin^{n-2i+1}\theta dF_{n-i(2n-2i-r)} dF'_{n-i(2n-2i-r)} dF_{2n-2i-r} & (2i \leq 2n-r) \ . \end{cases} \tag{4.26}$$

Alternatively, $M_i(Y)$ may be estimated by taking an IR projection onto an r-subspace $(0 \leq i < r < n)$. It is assumed that Y is opaque and the remainder of the specimen sufficiently transparent. By putting $q = i$, multiplying both sides of (2.4) by $V_i(Y|L_i)$, and then integrating over all pairs of subspaces L_i, L_r with $L_i \subset L_r$,

$$\int dL_i^r \int M_i^r(Y|L_r) dL_r = \int dL_{r(i)} \int dL_i \, M_i(Y) \ . \tag{4.27}$$

Hence $M_i^r(Y|L_r)$ is an unbiased estimator of $M_i(Y)$ with respect to an IR r-subspace L_r.

To calculate the variance of this projection estimator, we need the subspace analogue of Theorems 4.1 and 4.2, namely

$$dL_i^r dL_i'^r dL_r = \begin{cases} \sin^{r-n}\theta dL_{r(2i)} dL_i dL_i' & (2i \leq r) \\ \sin^{2i-n}\theta dL_{i(2i-r)} dL'_{i(2i-r)} dL_{2i-r} & (2i \geq r) \ . \end{cases} \tag{4.28}$$

Multiplying both sides by $V_i V_i' = V_i(Y|L_i) V_i(Y|L_i')$, integrating over all configurations and rearranging, we obtain

$$E\left[M_i^r(Y|L_r)\right]^2$$

$$= \frac{1}{\int dL_r\left(\int dL_i^r\right)^2} \times \begin{cases} \iint dL^n_{r(2i)} \iint V_i V_i' \sin^{r-n}\theta dL_i dL_i' & (2i \le r) \\ \iiint V_i V_i' \sin^{2i-n}\theta dL_{i(2i-r)} dL'_{i(2i-r)} dL_{2i-r} & (2i \ge r) \end{cases} \quad (4.29)$$

from which the variance follows easily.

As an example, take $n = 3$, $r = 2$, $i = 1$, and suppose that Y is compact convex. Then $M_1(Y)$ is the mean caliper diameter M of Y and $M_1^2(Y|L_2)$ is π^{-1} times the perimeter length B of $Y|L_2$. From (4.29),

$$\mathrm{Var}(B) = \frac{1}{2\pi} \iint L(Y|L_1)L(Y|L_1')\sin^{-1}\theta dL_1 dL_1' - \pi^2 M^2 . \quad (4.30)$$

4.6. Fractals and varieties

Let Y be a Borel set having finite m-dimensional integral geometric measure $I^m(Y)$. From (3.35),

$$\frac{\sigma_{m+1}\sigma_{n-m+1}}{2\sigma_{n+1}} M_{n-r}(X) I^{m+r-n}(Y \cap F_r) \quad (r \ge n-m \ge 0) \quad (4.31)$$

is an unbiased estimator of $I^m(Y)$ with respect to an IUR r-flat hitting X . Its variance is

$$-[I^m(Y)]^2 + \left(\frac{\sigma_{m+1}\sigma_{n-m+1}}{\sigma_{n+1}}\right)^2 \frac{M_{n-r}(X)}{\int dL_r\left(\int dL^r_{n-m}\right)^2}$$

$$\times \begin{cases} \iint dL_{r(2n-2m+1)} \iint PP'\delta^{r-n} \sin^{r-n}\theta dF_{n-m} dF'_{n-m} & (2m > 2n-r) \\ \iiint PP' \sin^{n-2m+1}\theta dF_{n-m(2n-2m-r)} dF'_{n-m(2n-2m-r)} dF_{2n-2m-r} & (2m \le 2n-r) \end{cases} \quad (4.32)$$

where P, P' are the numbers of intersections between Y and F_{n-m}, F'_{n-m}

respectively.

Quadrat estimation may also be used - we omit the details.

4.7. Functions and vector fields

With smoothness conditions on f, $\mathbf{f}$ as specified in Section 3.7, we obtain from the results in that section that

$$M_{n-r}(X) \int_{X\cap F_r} \left\{ \begin{array}{c} f \\ n\mathbf{f}_r/r \\ \sigma_n\omega_{r-1}|\mathbf{f}_r|/\sigma_r\omega_{n-1} \\ n\nabla_r f/r \\ n\nabla_r.\mathbf{f}_r/r \\ n(n-1)\nabla_r\times\mathbf{f}_r/r(r-1) \\ n\nabla_r^2 f/r \end{array} \right\} d\mathrm{x}_r$$

is an unbiased estimator of

$$\int_X \left\{ \begin{array}{c} f \\ \mathbf{f} \\ |\mathbf{f}| \\ \nabla f \\ \nabla.\mathbf{f} \\ \nabla\times\mathbf{f} \\ \nabla^2 f \end{array} \right\} d\mathrm{x} \; . \qquad (4.33)$$

The variance of the first of these estimators, for example, is

$$\frac{\sigma_r}{\sigma_n} M_{n-r}(X) \int_X \int_X f(\mathrm{x})f(\mathrm{y})|\mathrm{x}-\mathrm{y}|^{r-n} d\mathrm{x}d\mathrm{y} - \left(\int_X f(\mathrm{x})d\mathrm{x}\right)^2 . \qquad (4.34)$$

CHAPTER 5

RATIO ESTIMATION AND WEIGHTED SAMPLING

5.1. Introduction

The sectional and quadrat estimators of the preceding chapter have a common defect - they ignore the specimen cross-section $X \cap F_r$ or intersection $X \cap Q$ and concentrate on the feature trace $Y \cap F_r$ or $Y \cap Q$. The variance may be due largely to variation in the extent of the sample. In classical sampling theory, ratio estimation is frequently used to overcome this problem, but unless used in conjunction with "probability proportional to size" sampling, bias is introduced. It happens that, in our geometrical setting, this latter type of sampling is fairly easy to implement and reduces the chance of obtaining situations where F_r (or Q) only grazes the specimen. Furthermore, under certain assumptions of homogeneity of structure, weighted sampling reduces mean square error in addition to eliminating bias. In the context of weighted sampling, it is possible to give a rigorous interpretation of the fundamental formulae of stereology. Although these formulae have been widely used by practising stereologists, the only rigorous derivations of them given to date have required the assumption of either

(i) large sample sizes and independent observations (Mayhew and Cruz-Orive (1974), Miles (1972a), or

(ii) infinite specimens (Giger (1970)), or

(iii) random structures and infinite specimens (Miles (1976), Serra (1969)).

Our intrepretation is valid for either single or multiple observations made on an arbitrary deterministic specimen. The work in this chapter stems largely from Davy and Miles (1977) and Miles and Davy (1976a, 1976b).

5.2. Ratio estimators

One of the best known (and historically the first) of the fundamental formulae of stereology is the so-called Delesse principle (Section 1.1), which states that the proportion of area (A_A) covered by a particular phase of material on a planar cross-section is equal to the proportion of volume (V_V) occupied by that phase in three dimensions. In terms of our model, the natural way to define A_A is as $A(Y \cap F_2)/A(X \cap F_2)$, and its expectation must be evaluated with respect to a random plane having a certain probability distribution. The truth of the assertion "$E(A_A) = V_V$" depends upon the particular randomization chosen. For an IUR plane through X (letting A denote the cross-sectional area $A(X \cap F_2)$),

$$\begin{aligned} E(A_A) &= E[A(Y \cap F_2)]/EA - E[A(Y \cap F_2)]/EA + E(A_A)EA/EA \\ &= V(Y)/V(X) - \mathrm{Cov}(A_A, A)/EA \\ &= V_V - \mathrm{Cov}(A_A, A/EA) . \end{aligned} \tag{5.1}$$

Hence A_A is a *biased* estimator with respect to an IUR plane, the bias being $-\mathrm{Cov}(A_A, A/EA)$. The bias may be positive or negative, depending upon the spatial distribution of Y within X. For example, in the extreme cases depicted in Figure 5.1, the covariance between A_A and A is positive in (a) and negative in (b). That is, the bias is negative and positive when the spatial distribution of Y within X is respectively central and peripheral.

However there does exist a probability distribution over the planes through X such that A_A has expectation V_V for arbitrary compact X, Y. This distribution is obtained by weighting the IUR distribution according

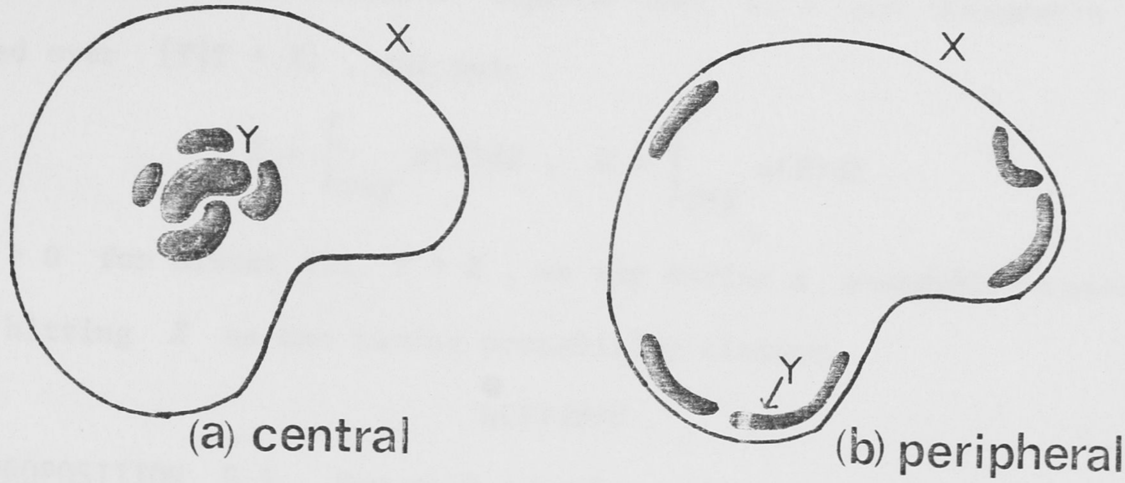


FIGURE 5.1. Two extreme cases of the possible spatial distributions of the embedded feature Y within the specimen X.

to the area of intersection $A(X \cap F_2)$, i.e. the probability element is

$$A(X \cap F_2)dF_2/\int A(X \cap F_2)dF_2 = A(X \cap F_2)dF_2/2\pi V(X) . \qquad (5.2)$$

In this way more importance is placed upon the cross-sections of larger extent and small probability is attached to sections which only graze the specimen and hence contain little information. To see that the bias is indeed eliminated, observe that

$$E_A(A_A) = \int \frac{A(Y\cap F_2)}{A(X\cap F_2)} A(X \cap F_2)dF_2/2\pi V(X)$$

$$= V(Y)/V(X) = V_V . \qquad (5.3)$$

Here the subscript A on the expectation operator signifies that an area-weighted plane has been used. A possible objection to the use of area-weighted sectioning could be difficulty of implementation. However in Section 2.3 it was shown that such a plane can be generated by first choosing a UR point of X and then (independently) constructing an IR plane containing this point.

We shall now generalize this example.

Let T (the probe) denote either an r-flat F_r $(0 \leq r < n)$ or an

r-dimensional quadrat Q $(0 < r \leq n)$, and let dT be the element of the appropriate invariant measure. Suppose that z, u are integrable functions defined over $\{T|T \uparrow X\}$, and put

$$Z = \int_{T\uparrow X} z(T)dT \ , \quad U = \int_{T\uparrow X} u(T)dT \ . \tag{5.4}$$

If $u > 0$ for almost all $T \uparrow X$, we may define a *u-weighted* random probe hitting X as one having probability element

$$u(T)dT/U \ . \tag{5.5}$$

PROPOSITION 5.1. *For such a probe,*

$$E_u(z/u) = Z/U \ , \tag{5.6}$$

where the subscript u *refers to expectation with respect to* u*-weighting. (The absence of a subscript will henceforth mean no weighting.) In other words,* z/u *is an unbiased estimator of* Z/U *with respect to a* u*-weighted probe.*

Proof.

$$E_u(z/u) = \int_{T\uparrow X} \frac{z}{u} \, . \, udT/U = Z/U \ . \tag{5.7}$$

In most applications, U is known, so that (5.6) enables estimation of the quantity Z .

PROPOSITION 5.2. *The bias in using the ratio* z/u *with respect to a* u'*-weighted probe is* $\int_{T\uparrow X} dT.\text{Cov}\left(\frac{u'}{U'} - \frac{u}{U}, \frac{z}{u}\right)$.

Proof.

$$E_{u'}\left(\frac{z}{u}\right) - Z/U = E\left(\frac{zu'}{u}\right)/Eu' - \int_{T\uparrow X} dT.E\left(\frac{zu}{u}\right)/U$$

$$= \int_{T\uparrow X} dT.\text{Cov}\left(\frac{u'}{U'} - \frac{u}{U}, \frac{z}{u}\right) \ . \tag{5.8}$$

In particular, putting $u' \equiv 1$, the bias when an IUR probe is used is $- \int_{T\uparrow X} dT.\text{Cov}(u/U, z/u)$.

In Chapter 3, we presented many equations of the type (5.4) and hence there is a large number of particular examples of (5.6). Putting $n = 3$, $r = 2$, $T = F_2$, $z(F_2) = A(Y \cap F_2)$, $u(F_2) = A(X \cap F_2)$ yields the Delesse formula (5.3). Putting $n = 3$, $r = 1$, T = line segment Q, $z(Q) = \int_{X \cap Q} f dx$, $u(Q) = L(X \cap Q)$ yields

$$E_L\left(\frac{1}{L(X \cap Q)} \int_{X \cap Q} f dx\right) = \int_X f d\mathrm{x}/V(X) . \tag{5.9}$$

z need not be real-valued. If we put $\mathrm{z}(F_r) = \mathrm{q}_{n+i-r}(X \cap F_r)$, $u(F_r) = W_{n+i-r}(X \cap F_r)$, we obtain from (3.9) and (3.13),

$$E_u\left[\frac{\mathrm{q}_{n+i-r}(X \cap F_r)}{W_{n+i-r}(X \cap F_r)}\right] = \frac{\mathrm{q}_i(X)}{W_i(X)} ,$$

i.e.

$$E_u\left[\mathrm{p}_i^r(X \cap F_r)\right] = \mathrm{p}_i(X) . \tag{5.10}$$

For $i = 0$, (5.10) states that the centroid of a V_r-weighted cross-section of X is an unbiased estimator of the centroid of X.

5.3. Comparisons of mean square error

The first question which we shall consider is whether or not the elimination of bias in ratio estimators via the use of weighted sampling is bought at a cost of increased mean square error (MSE). (The mean squared distance from the property to be estimated is a more appropriate measure of imprecision than the variance in the case of biased estimators.) It is easy to prove the following.

PROPOSITION 5.3. *The difference in* MSE *for* z/u *between a* u'-*weighted and an* IUR *probe is given by*

$$\mathrm{MSE}_{u'}(z/u) - \mathrm{MSE}(z/u) = \mathrm{Cov}\left[u'/U', \frac{z}{u}\left(\frac{z}{u} - 2\frac{Z}{U}\right)\right] . \tag{5.11}$$

In particular, putting $u' = u$,

$$\mathrm{Var}_u(z/u) - \mathrm{MSE}(z/u) = \mathrm{Cov}\left[u/U, \frac{z}{u}\left(\frac{z}{u} - 2\frac{Z}{U}\right)\right] . \tag{5.12}$$

Let us examine the covariance expression in (5.12) more closely. The second term within the square brackets is a quadratic in z/u, with minimum at $z/u = Z/U$. If u is a measure of the sample size (e.g. $A(X \cap F_2)$), we should expect that in a homogeneous structure, the deviations of the estimator z/u from Z/U will tend to be smaller for larger u. The covariance would then be negative, meaning that the variance of the weighted estimator is smaller than the MSE of the biased IUR estimator.

Summarizing, weighted probes are preferable (on the basis of mean square error) provided that the sample is sufficiently homogeneous to ensure that $\mathrm{Cov}\left[u/U,\ \frac{z}{u}\left(\frac{z}{u} - 2\,\frac{Z}{U}\right)\right]$ is negative.

Next we ask whether or not ratio estimation is preferable to the unbiased non-ratio procedures described in Chapter 4. Within that chapter are many instances of estimation of the form

$$E\alpha = \int_{T\uparrow X} \alpha(T)dT \Big/ \int_{T\uparrow X} dT = Z \tag{5.13}$$

with respect to an IUR or FUR probe T through X.

Put $z(T) = \alpha(T) / \int_{T\uparrow X} dT$ and consider the ratio estimator Uz/u of Z with respect to a u-weighted probe.

PROPOSITION 5.4.

$$\mathrm{Var}(\alpha) - \mathrm{Var}_u(Uz/u) = \mathrm{Cov}\left(u,\ \alpha^2/u\right) . \tag{5.14}$$

Thus the optimal weighting factor u is that which maximizes $\mathrm{Cov}\left(u, \alpha^2/u\right)$. If in fact $u \equiv \alpha$, $\mathrm{Cov}\left(u, \alpha^2/u\right) = \mathrm{Cov}(\alpha, \alpha) = \mathrm{Var}(\alpha)$, leading to the tautological statement that if we know what we are trying to estimate, then we can estimate it with perfect precision!

When u is approximately constant (as, for example, is the case when T is a quadrat whose mean caliper diameter is very small compared to that of the specimen X and $u = V_r(X \cap Q)$), then the covariance is close to zero and there is little difference between the variances of the two estimators. In this case $\mathrm{Var}_u(Uz/u)$ may be approximated by the appropriate variance formula of Chapter 4.

Another expression, although not very useful, for the variance is

$$\mathrm{Var}_u(Uz/u) = \int_{T\uparrow X} dT.ZU\ \mathrm{Cov}\left(\frac{z}{Z} - \frac{u}{U}, \frac{z}{u}\right) . \tag{5.15}$$

The final comparison made within this section is between flats of differing dimension.

Suppose as before that $Z = \int z(F_r)dF_r$. (The range of integration $T \uparrow X$ is omitted for simplicity - it is understood that $z(T) = 0$ for $T \not\uparrow X$.) By (2.9), Z may also be written as

$$\begin{aligned} Z &= \int \left(\int z(F_r^s)dF_r^s\right)dF_s / \int dF_{s(r)} \\ &= \int z'(F_s)dF_s \quad (0 \le r < s < n) \end{aligned} \tag{5.16}$$

where $z'(F_s) = \int z(F_r^s)dF_r^s / \int dF_{s(r)}$. Similarly, we may write $U = \int u'(F_s)dF_s$.

Hence z/u (with respect to a u-weighted r-flat) and z'/u' (with respect to a u'-weighted s-flat) are both unbiased estimates of Z/U . Which has the lower variance? Intuitively, one feels that the latter estimator, being based on a higher dimensional sample, would be preferable. This can indeed be shown to be true as follows.

From (2.9) and the definition of u' ,

$$\frac{u'}{U}\left(\frac{u\,dF_r^s}{\int u\,dF_r^s}\right)dF_s = \frac{u}{U}\left(\frac{dF_{s(r)}}{\int dF_{s(r)}}\right)dF_r . \tag{5.17}$$

This can be interpreted as saying that a u-weighted r-flat may be generated by a two-stage procedure - first a u'-weighted s-flat is generated and then a u-weighted r-flat chosen within this section. Hence we may use a conditional argument

$$E_u\left(\frac{z}{u}\middle|F_s\right) = \int z\left(F_r^s\right)dF_r^s \Big/ \int u\left(F_r^s\right)dF_r^s = z'/u' \ . \tag{5.18}$$

By Jensen's Inequality,

$$E_u\left[\left(\frac{z}{u}\right)^2\right] = E_{u'}\left[E_u\left\{\left(\frac{z}{u}\right)^2\middle|F_s\right\}\right] \geq E_{u'}\left[\left(\frac{z'}{u'}\right)^2\right] , \tag{5.19}$$

$$\Rightarrow \mathrm{Var}_u(z/u) \geq \mathrm{Var}_{u'}(z'/u') \ . \tag{5.20}$$

(Strict inequality holds provided that z'/u' is not a.e. equal to z/u .) Constrast this with the non-ratio estimation of the previous chapter, where it was shown that higher dimensional sections do not necessarily provide more efficient estimators.

As an example, consider the case where Y is a 3-dimensional sphere of radius s contained in a concentric sphere X of radius t . Letting L_L, A_A and V_V denote the proportions of length, area and volume occupied by Y (see Section 5.4), and using elementary integration, we may show that

$$\mathrm{Var}_L\left(L_L\right)/\left(V_V\right)^2 = 3(t/s)^4 - 2(t/s)^6 + 2(t/s)^3\left[(t/s)^2-1\right]^{3/2} - 1 \tag{5.21}$$

and

$$\mathrm{Var}_A\left(A_A\right)/\left(V_V\right)^2 = \tfrac{5}{2}(t/s)^3 - \tfrac{3}{2}(t/s)^5 + \tfrac{3}{4}(t/s)^2\left[(t/s)^2-1\right]^2 \ln\left(\frac{t/s+1}{t/s-1}\right) - 1 . \tag{5.22}$$

Table 5.1 lists the above relative variances for various values of t/s , together with relative variances for the non-ratio volume estimators of the previous chapter (Figure 4.1).

t/s	A_A	L_L	α_0	α_1	α_2
1	0	0	0	.125	.2
1.1	.0336	.1053	.331	.361	.32
1.2	.0889	.2575	.728	.62	.44
1.5	.2904	.8397	2.375	1.531	.8
2	.6603	2.1384	7	3.5	1.4

TABLE 5.1. Relative variances of volume estimators for embedded sphere

Note that, as predicted by (5.20), $\mathrm{Var}_A(A_A) < \mathrm{Var}_L(L_L) < \mathrm{Var}(\alpha_0/V(X))$. Also the ratio estimators A_A and L_L are uniformly more efficient than the corresponding non-ratio estimators α_2 and α_1 of the same dimension.

A similar comparison can be made between estimators based on r-flats and those based on r-dimensional quadrats.

The preceding analysis requires only minor modifications when we consider a vector-valued function $\mathrm{z}(T)$. Define $\mathrm{Z} = \int \mathrm{z}(T)dT$ and $\mathrm{MSE}_u(\mathrm{z}/u) = E_u|\mathrm{z}/u - \mathrm{Z}/U|^2$. Then

$$\mathrm{MSE}_{u'}(\mathrm{z}/u) - \mathrm{MSE}(\mathrm{z}/u) = \mathrm{Cov}[u'/U', \mathrm{z}.(\mathrm{z}/u - 2\mathrm{Z}/U)/u] , \tag{5.11'}$$

$$\mathrm{MSE}(\boldsymbol{\alpha}) - \mathrm{MSE}_u(U\mathrm{z}/u) = \mathrm{Cov}\left(u, |\boldsymbol{\alpha}|^2/u\right) \tag{5.14'}$$

where $\boldsymbol{\alpha}(T) = \int_{T\uparrow X} dT.\mathrm{z}(T)$ and

$$\mathrm{MSE}_u(\mathrm{z}/u) \geq \mathrm{MSE}_{u'}(\mathrm{z}'/u') \tag{5.18'}$$

where $\mathrm{z}'(F_s) = \int \mathrm{z}(F_r^s)dF_r^s / \int dF_{s(r)}$ $(0 \leq r < s < n)$.

5.4. The fundamental formulae of stereology

In this section it is shown how the ratio estimation theory presented above provides a rigorous derivation for all of the known fundamental formulae of stereology as well as furnishing further formulae with potential application. The usual notation for a ratio z/u in the stereological literature is z_u . Thus equation (5.6) may be rewritten as

$$E_u(z_u) = Z_U . \tag{5.23}$$

Usually the subscripts refer to specimen properties and the main letters to feature properties.

The following notation (chosen to conform closely to standard stereological usage) will be adopted.

3-dimensional quantities:

V volume

S surface

H mean areal projection (= $S/4$ when set is convex)

K integral of mean curvature $\frac{1}{2}\int(\kappa_1+\kappa_2)\,dS$ over oriented surface

M mean lineal projection or caliper diameter (= $K/2\pi$ when set is convex)

L length of space curve

2-dimensional quantities:

A area

B length of perimeter or of plane curve

M mean caliper diameter (= B/π when set is convex)

C total curvature (= $\pm 2\pi$ for a simple closed curve)

P number of points

I indicator function

1-dimensional quantities:

L length

P number of points

I indicator function

The following tables list Z_U for particular cases of z and u. It is supposed that u-weighting has been used.

$z \backslash u$	A	B	M	C	I	P
A	V_V	$\frac{4}{\pi} V_S$	V_H	V_K	V_M	-
B	$\frac{\pi}{4} S_V$	S_S	$\frac{\pi}{4} S_H$	$\frac{\pi}{4} S_K$	$\frac{\pi}{4} S_M$	-
M	H_V	$\frac{4}{\pi} H_S$	H_H	H_K	H_M	-
C	K_V	$\frac{4}{\pi} K_S$	K_H	K_K	K_M	-
I	M_V	$\frac{4}{\pi} M_S$	M_H	M_K	M_M	-
P	$\frac{1}{2}L_V$	$\frac{2}{\pi} L_S$	$\frac{1}{2}L_H$	$\frac{1}{2}L_K$	$\frac{1}{2}L_M$	L_L

TABLE 5.2. Planar section of 3-dimensional specimen

NOTES. (a) For a convex specimen set columns 2 and 3 and columns 4 and 5 can be seen to be equivalent when it is remembered that weighting according to u is equivalent to weighting according to a constant multiple of u.

(b) The I column gives the simple unweighted estimators of Chapter 4.

(c) The restriction of the table to the A, B, C and P rows and columns remains valid for planar quadrats Q provided that we measure only the internal quantities $B(\partial Y \cap Q)$, $C(\partial Y \cap Q)$, $B(\partial X \cap Q)$, $C(\partial X \cap Q)$ as described in Section 4.2.

(d) $E_A(A_A) = V_V$ remains valid for FUR sections or quadrats.

$z \backslash u$	L	I	P
L	$A_A(V_V)$	$A_M(V_H)$	$\frac{\pi}{2} A_B(2V_S)$
I	$M_A(H_V)$	$M_M(H_H)$	$\frac{\pi}{2} M_B(2H_S)$
P	$\frac{2}{\pi} B_A(\frac{1}{2}S_V)$	$\frac{2}{\pi} B_M(\frac{1}{2}S_H)$	$B_B(S_S)$

TABLE 5.3. Linear transect through 2- (3-) dimensional specimen

NOTES. (a) For a convex specimen set columns 2 and 3 are equivalent.

(b) The I column gives unweighted estimators.

(c) The table restricted to the L and P rows and columns remains valid for quadrats, provided that we measure $P(\partial Y \cap Q)$, $P(\partial X \cap Q)$.

(d) $E_L(L_L) = A_A$ or V_V remains valid for FUR rather than IUR sampling.

CHAPTER 6

NON-FLAT SECTIONS

6.1. Introduction

In this chapter we shall restrict ourselves, for ease of exposition, to E^3, and consider estimation based on certain types of non-flat 2-dimensional sections. The first of these is a wedge section, i.e. two half planes at angle ω $(0 < \omega < 2\pi)$ emanating from a common line. Such a section arises, for example, when we examine the two cut surfaces of a wedge of cake, or when we look at the cut surfaces of one of the four pieces obtained when a convex specimen is divided by two intersecting planes. The second type of section is a smooth curved surface, which could be obtained, for example, by polishing a metal specimen to a spherical shape.

It has long been conjectured that the number of particles per unit volume or, more generally, the integral of Gaussian curvature per unit volume cannot be estimated from two-dimensional data without further assumptions concerning the structure of the feature (e.g. that Y consists of non-overlapping spherical or ellipsoidal particles - see Wicksell (1925, 1926)). We shall see that estimation of these quantities is possible for quite general structures when using wedge sections, although difficulties arise when the wedge angle ω is too close to 0 or π. When using smoothly curved sections we can estimate the integrals of a variety of functions of curvature, although Gaussian curvature still seems to be inaccessible. The material in Section 6.2 is to appear in a joint paper with R.E. Miles (1978).

6.2. Wedge sections

Figure 6.1 illustrates a wedge section intersecting a feature set Y having smooth (twice continuously differentiable) surface. The wedge edge meets the surface at a point Q, having outward normal $\mathbf{N}$. The two faces a and b of the wedge intersect the surface in curves QP and QR. The tangent vectors to these curves through Q are denoted by $\mathbf{t}_a$ and $\mathbf{t}_b$ respectively. The curvatures along QP and QR at Q are κ_a and κ_b. $\mathbf{N}_a$, $\mathbf{N}_b$ are the normals to faces a and b, pointing outwards from the wedge. ψ_a, ψ_b are the angles between $\mathbf{N}_a$, $\mathbf{N}_b$ and $\mathbf{N}$, and ν is the angle between $\mathbf{t}_a$ and $\mathbf{t}_b$.

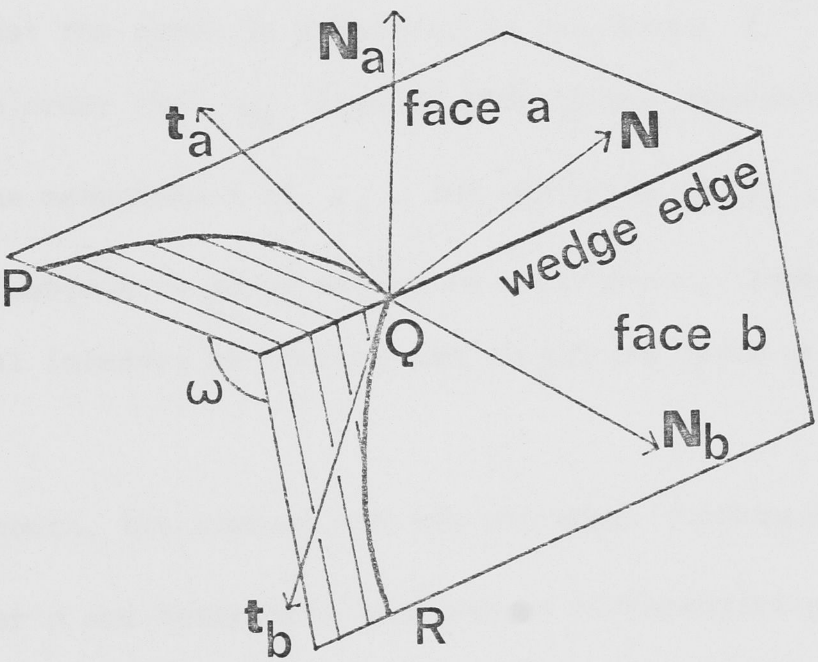


FIGURE 6.1. A wedge section intersecting the feature set Y

κ_a, κ_b, ψ_a, ψ_b and ν are all observable. If ρ_a, ρ_b are the angles between $\mathbf{t}_a$, $\mathbf{t}_b$ and the wedge edge (directed outwards from Y), then the following formulae from spherical trigonometry (Todhunter and Leathem (1963, Chapter III) enable determination of ψ_a, ψ_b and ν in terms of ρ_a, ρ_b and the wedge angle ω :

$$\cos \nu = \cos \rho_a \cos \rho_b + \sin \rho_a \sin \rho_b \cos \omega , \tag{6.1}$$

$$\cos \psi_a = \left(\cos \rho_a \sin \rho_b - \sin \rho_a \cos \rho_b \cos \omega\right)/\sin \nu , \tag{6.2}$$

$$\cos \psi_b = \left(\sin \rho_a \cos \rho_b - \cos \rho_a \sin \rho_b \cos \omega\right)/\sin \nu . \tag{6.3}$$

κ_a is the rate of change of tangent direction to QP with respect to arc length. Its sign is positive if the tangent is rotated towards Y and negative otherwise. One method of measuring κ_a is to construct a tangent to QP at arc length δ from Q and measure the signed angle ε between this tangent and $\mathbf{t}_a$. The appropriate estimate of κ_a is ε/δ. If the two constructed normals are independently distributed about their true directions with variance σ^2, then the variance of ε/δ is $2\sigma^2/\delta^2$. This means that the error is magnified by the factor δ^{-2}, but δ must be kept small in order that κ_a remains effectively constant over this interval. The measurement of κ_a, and similarly of κ_b, is therefore likely to be subject to large errors in measurement. Nevertheless it is of theoretical interest to investigate estimators based on $\kappa_a, \kappa_b, \psi_a, \psi_b$ and ν.

Being smooth, the surface has two principal curvatures κ_1, κ_2 at Q. The Euler's and Meusnier's theorems of differential geometry (Struik (1961, pp. 76, 81)) may be applied at Q for each of the wedge faces, regarded as plane sections of the surface. They yield

$$\begin{aligned} \kappa_a \sin \psi_a &= \kappa_1 \cos^2\varphi_a + \kappa_2 \sin^2\varphi_a , \\ \kappa_b \sin \psi_b &= \kappa_1 \cos^2\varphi_b + \kappa_2 \sin^2\varphi_b , \end{aligned} \tag{6.4}$$

where φ_a, φ_b are the angles between $\mathbf{t}_a, \mathbf{t}_b$ and the principal direction of curvature corresponding to κ_1. It is supposed without loss of generality that $\nu = \varphi_a - \varphi_b$

As we have in effect three unknowns $(\kappa_1, \kappa_2$ and one of $\varphi_a, \varphi_b)$ but only two simultaneous equations, we cannot solve for κ_1, κ_2 exactly. However, a quadratic equation relating κ_1 and κ_2 may be derived as follows. From (6.4),

$$\alpha \equiv \kappa_a \sin\psi_a - \kappa_b \sin\psi_b = (\kappa_1-\kappa_2)\sin\nu\sin(2\varphi_a+\nu) , \tag{6.5}$$

$$\beta \equiv \kappa_a \sin\psi_a + \kappa_b \sin\psi_b = (\kappa_1+\kappa_2) + (\kappa_1-\kappa_2)\cos\nu\cos(2\varphi_a+\nu) . \tag{6.6}$$

Equating values for $(\kappa_1-\kappa_2)^2\cos^2(2\varphi_a+\nu)$ from (6.5) and (6.6), we get

$$\{\beta-(\kappa_1+\kappa_2)\}^2\sec^2\nu = (\kappa_1-\kappa_2)^2 - \alpha^2\mathrm{cosec}^2\nu , \tag{6.7}$$

a quadratic equation in κ_1, κ_2.

Suppose that the wedge is isotropically oriented relative to the specimen. Then, given that the wedge edge passes through Q (strictly speaking, a small surface area element containing Q), an event we denote by $\uparrow$, the conditional density of the polar co-ordinates of the edge direction is

$$f(\eta, \xi|\uparrow) = 2\sin\eta\cos\eta\,(2\pi)^{-1} \quad (0 \le \eta \le \pi/2,\ 0 \le \xi < 2\pi) , \tag{6.8}$$

i.e. η and ξ are conditionally independent, with ξ uniform. Moreover, conditional on any value of (η, ξ), the orientation γ of the wedge about its edge is isotropic. We shall now exploit the conditional uniformity of ξ given η and γ. It implies that the conditional distribution of φ_a (or φ_b) given ν, ψ_a and ψ_b is uniform on $[0, 2\pi]$.

Hence, by (6.5),

$$E\left\{\left|\frac{\alpha}{\sin\nu}\right|^k \middle| \uparrow, \nu, \psi_a, \psi_b\right\}$$

$$= |\kappa_1-\kappa_2|^k \int_0^{2\pi} \sin^k(2\varphi_a+\nu)\,d\varphi_a/2\pi$$

$$= \pi^{-\frac{1}{2}}\left\{\Gamma\left(\frac{k+1}{2}\right)/\Gamma\left(\frac{k}{2}+1\right)\right\}|\kappa_1-\kappa_2|^k \quad (k = 1, 2, \ldots) , \tag{6.9}$$

$$E(\beta|\uparrow,\ \nu,\ \psi_a,\ \psi_b) = \kappa_1 + \kappa_2 \tag{6.10}$$

and

$$E\left[\beta^2|\uparrow,\ \nu,\ \psi_a,\ \psi_b\right] = (\kappa_1+\kappa_2)^2 + \tfrac{1}{2}(\kappa_1-\kappa_2)^2\cos^2\nu\ . \tag{6.11}$$

Higher order conditional moments of β may also be derived with little difficulty.

Combining (6.9) for $k = 2$ with (6.11), we have for an arbitrary function $h(\nu)$,

$$E\{\beta^2+h(\nu)\alpha^2|\uparrow,\ \nu\} = (\kappa_1+\kappa_2)^2 + \tfrac{1}{2}(\kappa_1-\kappa_2)^2\{\cos^2\nu+h(\nu)\sin^2\nu\}\ . \tag{6.12}$$

Choosing $h(\nu)$ to make the last term in (6.12) equal to $-(\kappa_1-\kappa_2)^2$, we have

$$E(\tau|\uparrow) = \kappa_1\kappa_2 \tag{6.13}$$

where

$$\tau = (\beta/2)^2 - \left(\frac{2+\cos^2\nu}{\sin^2\nu}\right)(\alpha/2)^2\ . \tag{6.14}$$

Let us now examine the stereological implications of the preceding analysis. As previously, we suppose that Y is a subset of the compact set X , and generate an IUR position of the wedge hitting X (i.e. an IUR line through X is generated, and then the wedge is oriented uniformly randomly about this edge). The wedge edge hits the surface of Y in an a.s. finite number m of points, each having associated τ value τ_i $(1 \le i \le m)$. Consider the statistic $\sum_1^m \tau_i$. Then, letting $I(dS)$ equal dS or 0 according as the wedge edge hits or misses the surface element dS of Y ,

$$E\left\{\sum \tau_i\right\} = E\left[\int_{\partial Y} \tau I(dS)\right]$$

$$= \int_{\partial Y} E(\tau|\uparrow)Pr(\uparrow)$$

$$= \int_{\partial Y} \kappa_1\kappa_2 dS/2H(X)$$

$$= G(\partial Y)/2H(X) \ . \tag{6.15}$$

Recall that G , as defined in Chapter 3, is the integral of Gaussian curvature.

Alternatively, ratio estimation may be used. Suppose that the wedge is generated by first giving the edge an L-weighted random position, and then isotropically orienting the wedge about its edge. Then

$$E_L\left\{\sum \tau_i/L\right\} = E\left\{\sum \tau_i\right\}/EL = G(\partial Y)/2V(X) \ . \tag{6.16}$$

In the case when Y consists of a number N of particles without holes, then $G(\partial Y) = 4\pi N$, and (6.16) becomes

$$E_L\left\{\sum \tau_i/2\pi L\right\} = N_V \ . \tag{6.17}$$

Thus, stereological estimation of N_V on the basis of two dimensional information *is* possible, a fact which has long been conjectured to be false except under very special assumptions.

In a similar fashion, we may derive from (6.12) the more general formula

$$E_L\left\{\sum\left(\beta_i^2 - \frac{2\lambda+\cos^2\nu_i}{\sin^2\nu_i}\alpha_i^2\right)/L\right\} = \int_{\partial Y}\left\{\left(\kappa_1+\kappa_2\right)^2-\lambda\left(\kappa_1-\kappa_2\right)^2\right\}dS/2V(X) \ . \tag{6.18}$$

For example, putting $\lambda = 1$ yields (6.16), while $\lambda = 0$ enables estimation of the integral of squared mean curvature $\left(\kappa_1+\kappa_2\right)^2/4$, a quantity which we shall meet in the next section.

Finally, observe that (6.18) remains valid when the wedge angle ω is random rather than fixed, and therefore applies to the case when the wedge is formed by two independent IR planes containing an L-weighted secant of X .

6.3. Stereology on curved surfaces

In this section we suppose once again that Y is bounded by a smooth surface, but instead of a wedge section we take our probe to be a smooth, bounded, oriented surface Q . At each point x on ∂Y there exist principal curvatures κ_1, κ_2 and a unit normal vector N . Let e be a unit vector having the direction of principal curvature corresponding to K_1 . (If x is an umbilic point e is chosen in a continuous fashion from the unit tangent vectors through x .) Likewise, let $\kappa_1', \kappa_2', \mathrm{N}'$ and e' be defined at each point on Q .

For general positions of Q hitting ∂Y , the intersection set $\partial Y \cap Q$ consists of a smooth curve, which has at each point a tangent vector $\mathrm{t} = (\mathrm{N} \times \mathrm{N}')/\sin\theta$ and a curvature vector $\mathrm{k} = \dfrac{d\mathrm{t}}{ds}$, where θ is the angle between N and N' and ds is the element of arc length along $Y \cap Q$. k may be resolved in directions perpendicular and parallel to each surface, yielding the decompositions

$$\mathrm{k} = \mathrm{k}_N + \mathrm{k}_g = \mathrm{k}_N' + \mathrm{k}_g' , \tag{6.19}$$

where $\mathrm{k}_N, \mathrm{k}_N'$ are the normal curvature vectors with respect to ∂Y and Q , and $\mathrm{k}_g, \mathrm{k}_g'$ are the geodesic curvature vectors.

We shall evaluate certain integrals of the form

$$\int_{Q\uparrow\partial Y} \int_{Q\cap\partial Y} f ds dQ , \tag{6.20}$$

where f is a function of either k, k_g or k_g' .

The position of Q together with the position of a point x on the intersection curve $Q \cap \partial Y$ may be jointly parametrized by

$(\theta, u, v, \varphi, u', v', \varphi')$ where (u, v) (resp. (u', v')) are the surface co-ordinates of x relative to Y (resp. Q), and φ (resp. φ') is the anti-clockwise angle between t and e (resp. e′). The ranges of the angle parameters are $0 \leq \theta \leq \pi$, $0 \leq \varphi, \varphi' < 2\pi$.

Blaschke (1949, p. 122) has given an expression for the joint element $dsdQ$ which, in our context, can be wirtten as

$$dsdQ = \sin^2\theta d\theta d\varphi d\varphi' dS(u, v) dS'(u', v') . \tag{6.21}$$

Here $dS(u, v)$ and $dS'(u', v')$ are the elements of surface measure on Q and Y , and will henceforth be abbreviated to dS and dS' .

The next step involves expressing the curvature vector k in terms of the above parameters. According to Euler's theorem,

$$\mathrm{k}_N = \left\{\kappa_1 \cos^2\varphi + \kappa_2 \sin^2\varphi\right\}\mathrm{N} = h\mathrm{N} \text{ , say} \tag{6.22}$$

and $\mathrm{k}'_N = h'\mathrm{N}'$. Note that $\mathrm{k}.\mathrm{N} = h$, $\mathrm{k}.\mathrm{t} = 0$ and $\mathrm{N}.\mathrm{N}' = \cos\theta$.

Provided that the two normal vectors N, N′ are linearly independent (i.e. $\sin\theta \neq 0$), which is a.e. true, we may write k as a linear combination of the form

$$\mathrm{k} = a\mathrm{N} + a'\mathrm{N}' . \tag{6.23}$$

From (6.23) and the comments immediately after (6.22), we obtain

$$h = a + a' \cos\theta \text{ , } h' = a' + a\cos\theta . \tag{6.24}$$

Solving these simultaneous equations for a, a' yields

$$a = \frac{h-h'\cos\theta}{\sin^2\theta} \text{ , } a' = \frac{h'-h\cos\theta}{\sin^2\theta} . \tag{6.25}$$

Hence

$$\begin{aligned} |\mathrm{k}|^2 = ah + a'h' &= \left(h^2+h'^2-2hh'\cos\theta\right)/\sin^2\theta \\ &= \left(\mathrm{k}_N.\mathrm{k}_N+\mathrm{k}'_N.\mathrm{k}'_N-2\mathrm{k}_N.\mathrm{k}'_N\right)/\sin^2\theta \\ &= |\mathrm{k}_N-\mathrm{k}'_N|^2/\sin^2\theta . \end{aligned} \tag{6.26}$$

Using Pythagoras' theorem,

$$|k_g|^2 = |k|^2 - |k_N|^2 = (h' - h\cos\theta)^2/\sin^2\theta \ . \tag{6.27}$$

By combining (6.21), (6.26) and (6.27), we can evaluate integrals of the type (6.20). First, putting $f \equiv 1$,

$$\begin{aligned}\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y} ds dQ &= \int_{Q\uparrow\partial Y} L(Q\cap\partial Y)dQ \\ &= \int_Q\int_{\partial Y}\int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi} \sin^2\theta d\theta d\varphi d\varphi' dS dS' \\ &= 2\pi^3 SS' \end{aligned} \tag{6.28}$$

where S, S' are the total surface areas of ∂Y, Q respectively.

Similarly

$$\begin{aligned}&\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y} |k|^2 ds dQ \\ &= \int_Q\int_{\partial Y}\int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi} \frac{(h^2+h'^2-2hh'\cos\theta)}{\sin^2\theta} \sin^2\theta d\theta d\varphi d\varphi' dS dS' \\ &= \pi\int_Q\int_{\partial Y}\int_0^{2\pi}\int_0^{2\pi} \left(h^2+h'^2\right) d\varphi d\varphi' dS dS' \ .\end{aligned}$$

But

$$\begin{aligned}\int_0^{2\pi} h^2 d\varphi &= 4\int_0^{\pi/2}\left\{\kappa_1\cos^2\varphi + \kappa_2\sin^2\varphi\right\}d\varphi \\ &= 3\pi\left(\frac{\kappa_1+\kappa_2}{2}\right)^2 - \pi\kappa_1\kappa_2 \ .\end{aligned}$$

Therefore

$$\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y} |k|^2 ds dQ = 2\pi^3 S(R'-G') + 2\pi^3(3R-G) \tag{6.29}$$

where G, G' are integrals of Gaussian curvature and R, R' are integrals of squared méan curvature over ∂Y and Q . (6.29) was first derived by Chen (1975).

In the case of squared geodesic curvature,

$$\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y}|k_g|^2 ds dQ = \int_Q\int_{\partial Y}\int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi}(h'-h\cos\theta)^2 d\theta d\varphi d\varphi' dS dS'$$

$$= 2\pi^3 S(3R'-G') + \pi^3 S'(3R-G) \ . \qquad (6.30)$$

By considerations of symmetry,

$$\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y}|k_g'|^2 ds dQ = \pi^3 S(3R'-G') + 2\pi^3 S'(3R-G) \ . \qquad (6.31)$$

Provided that h, h' are a.e. non-zero,

$$\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y}|k| ds dQ = \int_Q\int_{\partial Y}\int_0^{2\pi}\int_0^{2\pi}\frac{|h+h'|^3-|h-h'|^3}{3hh'} d\varphi d\varphi' dS dS' \ . \qquad (6.32)$$

Geodesic curvature is usually given a sign. If we adopt the convention $k_g = (h'-h\cos\theta)/\sin\theta$, then

$$\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y}k_g ds dQ = 4\pi S\int_Q\int_0^{2\pi}h' d\varphi' dS'$$

$$= 4\pi S\int_Q\int_0^{2\pi}\left\{\kappa_1'\cos^2\varphi'+\kappa_2'\sin^2\varphi'\right\}d\varphi' dS'$$

$$= 8\pi^2 S K' \qquad (6.33)$$

where K' is the integral of mean curvature over Q . Similarly

$$\int_{Q\uparrow\partial Y}\int_{Q\cap\partial Y}k_g' ds dQ = 8\pi^2 S' K \ . \qquad (6.34)$$

It is readily seen that, unless all principal curvatures are zero almost everywhere, integrals involving $|k|^m$, $|k_g|^m$ or $|k_g'|^m$ for $m > 2$ are infinite.

To give the preceding identities a stereological interpretation, we can proceed in three different ways - namely we can give Q an IUR position hitting the specimen X , we can weight the position according to the surface area of intersection of Q with X , or we can weight according to the length of intersection with ∂Y .

S-weighting is achieved by choosing independent UR points of Q and

X , superimposing these points, and giving Q an IR orientation. L-weighting is achieved by choosing independent UR points of Q and ∂Y , superimposing, and generating a random orientation for Q having probability element $\sin^2\theta d\theta d\varphi d\varphi'/2\pi^3$ (a consequence of (6.21)).

Using familiar arguments, and putting $\beta = \int_{Q\cap\partial Y} |k|^2 ds$, we obtain

$$E(\beta) = [\pi S'(3R-G)+\pi S(3R'-G')]/4E[V(X-Q)] , \tag{6.35}$$

$$E_S(\beta_S) = \frac{\pi}{4}[3R_V-G_V] + \alpha'.S_V , \tag{6.36}$$

where

$$\alpha' = \frac{\pi(3R'-G')}{4S'}$$

and

$$E_L(\beta_L) = \alpha + \alpha' \quad \text{(from (6.29))}. \tag{6.37}$$

The factor α' appearing in (6.36) and (6.37) is a measure of how tightly "curved" Q is. If Q is a sphere of radius r , for example, then $\alpha' = \pi/2r^2 \to 0$ as $r \to \infty$.

Similarly, if $\gamma = \int_{Q\cap\partial Y} |k'_g|^2 ds$ and $\xi = \int_{Q\cap\partial Y} k'_g ds$,

$$E\gamma = [\pi S'(3R-G)+\tfrac{1}{2}\pi S(3R'-G')]/4E[V(X-Q)] , \tag{6.38}$$

$$E_S\gamma_S = \frac{\pi}{4}[3R_V-G_V] + \tfrac{1}{2}\alpha' S_V , \tag{6.39}$$

$$E_L\gamma_L = \alpha + \tfrac{1}{2}\alpha' , \tag{6.40}$$

and

$$E\xi = S'K/E[V(X-Q)] , \tag{6.41}$$

$$E_S[\xi_S] = K_V , \tag{6.42}$$

$$E_L[\xi_L] = K_S . \tag{6.43}$$

In the case when Q is planar, (6.42) reduces to the Cahn-DeHoff (1967) relation $E_A(C_A) = K_V$.

CHAPTER 7

RANDOM SETS

7.1. Introduction

Until now the spatial structures which we have considered have been deterministic, randomness being introduced only by the method of sampling. This has had the advantage that the stereological techniques described may be applied to quite general, possibly highly organized structures such as those occurring in biological materials. Stochastic models for spatial structure raise the same type of objections as the use of parametric, rather than non-parametric, methods in classical statistics. We can rarely be sure that the model adopted is adequate to describe the physical system. Also, non-repeatability is often inherent in spatial structures - it is hard to visualize a particular structure as a realization of an entire population of possible outcomes. In spite of these drawbacks, it is well worthwhile to follow in the steps of time series analysts and develop stochastic models for spatial processes. The recent development of random set theory by Kendall (1974) and Matheron (1975) enables this to be done in a mathematically elegant fashion. The advantages over the deterministic formulation are that the plausibility of various random mechanisms behind an existing structure may be assessed and that certain stereological procedures are simplified.

7.2. Definitions and basic properties

In order to talk about a random set in E^n it is necessary to choose a suitable σ-algebra on a class of sets (usually the closed sets F). This is done in the following rather natural manner. The events

$\{F \mid F \in \mathcal{F}, F \uparrow Q\}$, Q compact, are used to generate a σ-algebra on $\mathcal{F}$. A random closed set (RACS) is simply a measurable mapping $\omega \longmapsto Y(\omega)$ from a probability space into $\mathcal{F}$.

The sets $\{Y \uparrow Q\}$ are measurable by definition for Q compact. What is not so obvious, but can be proved by the use of capacity theory (Meyer (1966)) is that these sets are measurable, at least with respect to the completion of the underlying probability measure, for all Borel sets Q in E^n . The hitting probability $T_Y(Q) = Pr\{Y \uparrow Q\}$, regarded as a functional over the compact subsets of E^n , determines the RACS uniquely. Moreover, for a given functional T_Y , there exists an associated RACS iff T_Y is an alternating Choquet capacity of infinite order such that $0 \leq T_Y \leq 1$ and $T_Y(\emptyset) = 0$. (see Matheron (1975), Theorem 2.2.1.)

Definitions and brief discussions of some properties which may be possessed by a random set are given below.

Stationarity. Y is stationary iff the functional T_Y is invariant under translation. Stationarity of Y implies stationarity of the associated indicator function I_Y , regarded as a stochastic process on E^n , but the converse is not true. A simple counterexample is provided by taking Y to be a uniform random point in the interval $[0, 1]$ of E^1 . The stochastic process I_Y is equivalent to the process which is zero everywhere, and is therefore stationary. However $Pr\{Y \uparrow [0, 1]\} = 1$ and $Pr\{Y \uparrow [1, 2]\} = 0$, so that the random set Y is non-stationary.

Isotropy. Y is isotropic iff the functional T_Y is invariant under rotations. Isotropy of Y implies isotropy of the covariance function of I_Y , but not conversely.

Separability. The advent of image analysers has stressed the fact that in practice there is no such thing as a continuous image - we are limited to 0-1 values over a set of raster points (see Giger (1975)). A separable random set is such that all information concerning it can be obtained from a countable set of points. Formally, there exists a countable dense subset D of E^n such that almost surely $Y = \overline{Y \cap D}$.

Infinite divisibility. Y is said to be infinitely divisible if for any integer m , it is equivalent to the superposition (union) of m independent, identically distributed random sets Y_i . Union in random set theory plays a role analogous to addition in real variable theory.

Stability. Y is stable iff for any integer $m > 0$, there exists $\lambda_m > 0$ such that the union of m independent realizations of Y is equivalent to $\lambda_m Y$.

Semi-Markovian property. Markovian properties are usually dependent upon ordering (on the real line) or upon the concept of nearest neighbours (on a lattice). Neither of these are applicable to random sets. However there is a Markov type property which has been introduced into random set theory by Matheron (1975). Y is said to be semi-Markov iff

$$\left[1-T_Y(Q \cup Q' \cup Q'')\right]\left[1-T_Y(Q'')\right] = \left[1-T_Y(Q \cup Q'')\right]\left[1-T_Y(Q' \cup Q'')\right]$$

for any compact sets Q, Q' and Q'' such that Q'' separates Q and Q' (i.e. Q'' intercepts any line segment with one endpoint in Q and the other other in Q'). In other words, the RACS $Y \cap Q$ and $Y \cap Q'$ are independent conditional upon $Y \cap Q'' = \emptyset$.

7.3. Quermass densities

In Chapter 5 we gave an interpretation of the stereological ratios such as V_V, S_V in terms of a deterministic feature set contained within a compact specimen set. We now give a somewhat different interpretation in terms of random sets.

Suppose that Y is a stationary, isotropic RACS such that for any $Q \in K$, $Y \cap Q \in K$ a.s. Assume further that the quermassintegral $W_i(Y \cap Q)$ is a random variable whose expectation, regarded as a functional of Q, is finite and continuous over $C/\{\emptyset\}$. Note that

(i) $E[W_i(Y \cap \cdot)]$ is additive, due to the additivity of W_i and linearity of the expectation operator, and

(ii) $E[W_i(Y \cap \cdot)]$ is invariant under Euclidean motions, due to the stationarity and isotropy of Y.

Hence, by Theorem IV of Hadwiger (1957, Chapter 6), there exist constants c_{ij} such that

$$E[W_i(Y \cap Q)] = \sum_{j=0}^{n} c_{ij} W_j(Q) \quad (Q \in K,\ 0 \leq i \leq n) \ , \tag{7.1}$$

PROPOSITION 7.1.

$$c_{ij} = \begin{cases} \dfrac{\binom{i}{i-j}\omega_{n+j-i}\,\omega_i\,\omega_{n-j}}{\binom{n}{n-j}\omega_{n-i}\,\omega_{i-j}\,\omega_j\,\omega_n}\, c_{n,n+j-i} & (j \leq i) \\ \\ 0 & (j > i) \ . \end{cases} \tag{7.2}$$

Proof.

$$E[W_i(Y \cap Q)] = \frac{\omega_n}{\omega_{n-i}\int dL_i} E\left[\int \chi(Y \cap Q \cap F_i)\, dF_i\right] \quad \text{(by (3.8))}$$

$$= \frac{\omega_n}{\omega_{n-i}\int dL_i} \int E[\chi(Y \cap Q \cap F_i)]\, dF_i \quad \text{(Fubini's Theorem)}$$

$$= \sum_{j=0}^{n} \frac{c_{n,j}}{\omega_{n-i}\int dL_i} \int W_j(Q \cap F_i)\, dF_i \quad \text{(by (7.1))}$$

$$= \sum_{j=n-i}^{n} \frac{c_{n,j}\,\omega_j \binom{i}{n-j}}{\omega_{n-i}\int dL_i\,\omega_{i+j-n}\binom{n}{n-j}} \int W^i_{i+j-n}(Q \cap F_i)\, dF_i \quad \text{(by (3.9))}$$

$$= \sum_{j=0}^{i} \binom{i}{i-j} \frac{\omega_{n+j-i}\,\omega_i\,\omega_{n-j}}{\omega_{n-i}\,\omega_{i-j}\,\omega_j}\, c_{n,n+j-i} W_j(Q) \quad \text{(by (3.3))}\ .$$

The proposition now follows by equating coefficients. //

Let us now interpret the constants c_{ij} in terms of the random set Y .

PROPOSITION 7.2.

$$\lim_{t\to\infty} E[W_i(Y \cap tO)]/V(tO) = c_{n,n-i}/\binom{n}{i}\ . \tag{7.3}$$

Proof. From Proposition 7.1, the left-hand side is equal to

$$\lim_{t\to\infty} \sum_{j=0}^{i} \frac{\binom{i}{i-j}\omega_{n+j-i}\,\omega_i\,\omega_{n-j}}{\binom{n}{i-j}\omega_{n-i}\,\omega_{i-j}\,\omega_j}\, c_{n,n+j-i}\, t^{n-j}/\omega_n t^n = c_{n,n-i}/\binom{n}{i}\ . \qquad //$$

In fact Proposition 7.2 remains valid when tO is replaced by any increasing sequence of compact convex sets $Q(t)$ which satisfy the conditions $V[Q(t)] \to \infty$ and $W_i[Q(t)]/V[Q(t)] \to 0$ for $i > 0$. Hence if we put $D_i = c_{n,n-i}/\binom{n}{i}$, we may interpret D_i as the expected ith quermassintegral per unit volume. D_i will be called the ith quermass density. In standard stereological notation, the interpretations of D_i

for $n = 1, 2, 3$ are given by the following table.

$n \backslash i$	0	1	2	3
1	L_L	P_L		
2	A_A	$\frac{1}{2}B_A$	$\frac{1}{2}C_A$	
3	V_V	$\frac{1}{3}S_V$	$\frac{1}{3}K_V$	$\frac{1}{3}G_V$

TABLE 7.1. Interpretation of the quermass densities D_i

Often $W_i(Y \cap tO)/V(tO)$ converges almost surely as well as in mean, so that D_i can be regarded as an ergodic limit.

Equation (7.1) may be rewritten as

$$E\left[W_i(Y \cap Q)\right] = \frac{\omega_i}{\omega_n \omega_{n-i}} \sum_{j=0}^{i} \frac{\omega_{n-j}\omega_{n+j-i}}{\omega_j \omega_{i-j}} \binom{i}{j} D_{i-j} W_j(Q) \quad (0 \le i \le n) . \quad (7.4)$$

The fundamental formulae of stereology may be interpreted in terms of quermass densities as follows. For each $r < n$, Y determines an r-dimensional RACS obtained by intersecting Y with an arbitrary r-flat. The density D_i^r associated with this sectional RACS is related to D_i. Let Q be a flat r-dimensional set. Then

$$E[\chi(Q \cap Y)] = \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} D_{n-i} W_i(Q) \quad \left(\text{from } (7.4),\right.$$

$$\left.\text{remembering that } \chi = W_n/\omega_n\right)$$

$$= \sum_{i=n-r}^{n} \binom{r}{n-i} D_{n-i} \frac{\omega_i}{\omega_{r+i-n}} W^r_{r+i-n}(Q) \quad (\text{from } (3.9)).$$

Also,

$$E[\chi(Q \cap Y)] = \frac{1}{\omega_r} \sum_{i=0}^{r} \binom{r}{i} D_i^r W_i^r(Q) . \quad (7.5)$$

By equating coefficients of $W_i^r(Q)$, we obtain

PROPOSITION 7.3.

$$D_i^r = \frac{\omega_r \omega_{n-i}}{\omega_n \omega_{r-i}} D_i \quad (0 \le i \le r < n) . \qquad // \tag{7.6}$$

Particular examples include $P_P = L_L = A_A = V_V$ and $\pi P_L/4 = \frac{1}{2}B_A = \pi S_V/8$. Note the agreement with tables (5.2) and (5.3).

For the random open set Y^C , we define

$$D_i(Y^C) = \begin{cases} 1 - D_0(Y) & (i = 0) \\ (-1)^{i-1} D_0(Y) & (n \ge i > 0) . \end{cases} \tag{7.7}$$

Under this extended definition, table (7.1) and equation (7.6) remain valid.

7.4. Operations on random sets

Before going on to describe some particular models for random sets, let us examine some methods of building up more complicated models from simpler ones. If Y_1 and Y_2 are independent RACS, it may be shown by certain measurability results of Matheron (1975, pp. 7, 9, 19) that $Y_1 \cup Y_2$, $Y_1 \cap Y_2$ and (for Y_1 or Y_2 a.s. compact) $Y_1 + Y_2$ are also RACS. By the definition of T_Y ,

$$\begin{aligned} T_{Y_1 \cup Y_2}(Q) &= 1 - Pr\left[Q \cap (Y_1 \cup Y_2) = \emptyset\right] \\ &= 1 - Pr\left[(Q \cap Y_1) = \emptyset \text{ and } (Q \cap Y_2) = \emptyset\right] \\ &= 1 - (1 - T_{Y_1}(Q)).(1 - T_{Y_2}(Q)) \\ &= T_{Y_1}(Q) + T_{Y_2}(Q) - T_{Y_1}(Q).T_{Y_2}(Q) ; \end{aligned} \tag{7.8}$$

$$T_{Y_1 \cap Y_2}(Q) = Pr\left[Y_1 \cap Y_2 \cap Q \ne \emptyset\right] = E_{Y_2}\left[T_{Y_1}(Y_2 \cap Q)\right] = E_{Y_1}\left[T_{Y_2}(Y_1 \cap Q)\right] , \tag{7.9}$$

$$T_{Y_1 + Y_2}(Q) = Pr\left[(Y_1 + Y_2) \cap Q \ne \emptyset\right] = E_{Y_2}\left[T_{Y_1}(Q - Y_2)\right] = E_{Y_1}\left[T_{Y_2}(Q - Y_1)\right] . \tag{7.10}$$

In the case when Y_1 and Y_2 satisfy the conditions of the preceding section, we may also consider the quermass densities for $Y_1 \cup Y_2$ and $Y_1 \cap Y_2$.

PROPOSITION 7.4.

$$D_i(Y_1 \cup Y_2) = \sum_{k=0}^{i} \binom{i}{k} \frac{\omega_i \omega_{n-k} \omega_{n-i+k}}{\omega_n \omega_{n-i} \omega_k \omega_{i-k}} D_k(Y_1) D_{i-k}(Y_2) . \tag{7.11}$$

Proof. For $Q \in K$,

$E[\chi(Y_1 \cap (Y_2 \cap Q))]$

$$= \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} D_{n-i}(Y_1) E[W_i(Y_2 \cap Q)]$$

$$= \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} D_{n-i}(Y_1) \sum_{j=0}^{i} \binom{i}{j} \frac{\omega_i \omega_{n-j} \omega_{n+j-i}}{\omega_n \omega_{n-i} \omega_j \omega_{i-j}} D_{i-j}(Y_2) \times W_j(Q)$$

$$= \frac{1}{\omega_n} \sum_{j=0}^{n} \binom{n}{j} \left[\sum_{k=0}^{n-j} \binom{n-j}{k} \frac{\omega_{n-j} \omega_{n-k} \omega_{k+j}}{\omega_n \omega_k \omega_j \omega_{n-k-j}} D_k(Y_1) D_{n-k-j}(Y_2) \right] \times W_j(Q) . \tag{7.12}$$

Also,

$$E[\chi((Y_1 \cap Y_2) \cap Q)] = \frac{1}{\omega_n} \sum_{j=0}^{n} \binom{n}{j} D_{n-j}(Y_1 \cap Y_2) W_j(Q) . \tag{7.13}$$

By observing that the left-hand sides of (7.12) and (7.13) are equal, and by equating coefficients of $W_j(Q)$ on the right-hand sides, the proposition follows immediately. An example, taking $n = i = 3$, is

$$G_V(Y_1 \cap Y_2) = G_V(Y_1) V_V(Y_2) + K_V(Y_1) S_V(Y_2) + S_V(Y_1) K_V(Y_2) + V_V(Y_1) G_V(Y_2) .$$

COROLLARY 7.5. *The proposition* (7.4) *can be iterated, in the sense that if* $Y_1, \ldots, Y_m$ *are independent* RACS *satisfying the conditions of Section* 7.3,

$$D_i\left(\bigcap_{j=1}^{m} Y_j\right) = \sum_{k_1+\ldots+k_m=i} \frac{i!\,\omega_i \omega_n}{\omega_{n-i}} \prod_{j=1}^{m} \frac{\omega_{n-k_j}}{k_j!\,\omega_{k_j} \omega_n} D_{k_j}(Y_j) \quad (0 \le i \le n) . \tag{7.14}$$

COROLLARY 7.6.

$$D_i(Y_1 \cup Y_2) = D_i(Y_1) + D_i(Y_2) - \sum_{k=0}^{i} \binom{i}{k} \frac{\omega_i \omega_{n-k} \omega_{n-i+k}}{\omega_n \omega_{n-i} \omega_k \omega_{i-k}} D_k(Y_1) D_{i-k}(Y_2) \; . \quad (7.15)$$

Proof. From Proposition 7.2, D_i is additive, and hence (7.11) may be used to derive (7.15).

7.5. A Poisson family of models

An important family of RACS is defined by the functional

$$T_Y(Q) = 1 - \exp\left\{-\lambda E\left[V_q\left(G^q - Q^*\right)\right]\right\} \quad (Q \text{ compact}) \quad (0 \le q \le n) \; , \quad (7.16)$$

where λ is a constant, G^q is an a.s. compact convex RACS within the fixed subspace L_q, and Q^* is obtained by giving Q an IR rotation independent of G^q and then projecting it onto L_q. When $q = 0$, the random set is a.s. either $\emptyset$ or E^n.

The RACS Y having hitting functional (7.16) can be described (see Matheron (1975, p. 148)) as the union of the points of a process on the space of all convex cylinder sets with compact q-dimensional base (on C for $q = n$). The measure underlying the point process is invariant with respect to Euclidean motions. T_Y is the probability that at least one point (i.e. cylinder) of the process hits Q.

When G^q is a.s. a single point, Y is stable, and is in fact an $(n\text{-}q)$-dimensional Poisson flat process (see Miles (1971b)). When $q = n$, Y is the union of a Poisson process on C, and is called a Poisson grain process. The centroids of the constituent particles (the "grains") are distributed according to a homogeneous Poisson point process on E^n, and the grains are independent realizations of the a.s. compact convex isotropic RACS G^n. For $0 < q < n$, the RACS can best be visualized, if the reader will forgive the informality, as a system of interpenetrating pieces of infinite pasta. The cross-sections of the pasta are distributed according to the RACS G^q. Thus $q = n - 1$ corresponds to n-dimensional spaghetti (infinite straight tubes), and $q = 1$ to n-dimensional lasagne (hyperslabs).

It is clear both from (7.16) and from the point process representation that Y is stationary and isotropic. Furthermore, it is fairly easy to prove that Y is infinitely divisible and semi-Markov. The proofs are omitted, as they are to be found in Matheron (1975), who also proves a type of converse. A RACS is stationary, isotropic, infinitely divisible and semi-Markov iff it can be represented as the union of $n + 1$ independent RACS having functionals of the type (7.16), where q ranges from 0 to n.

For $Q \in C$, (7.16) may be written as

$$T_Y(Q) = 1 - \exp\left\{-\frac{\lambda}{\omega_n}\sum_{i=0}^{q}\binom{q}{i}W_{n-i}(Q)E\left[W_i^q(G^q)\right]\right\}$$

$$= 1 - \exp\left\{-\frac{1}{\omega_n}\sum_{i=0}^{n}\binom{n}{i}d_iW_{n-i}(Q)\right\} \qquad (7.17)$$

where

$$d_i = \begin{cases}\lambda\binom{q}{i}E\left[W_i^q(G^q)\right]/\binom{n}{i} & (0 \le i \le q)\\ 0 \quad \text{otherwise.}\end{cases}$$

Observe that $Y \cap Q \in K$ a.s. for $Q \in K$. Let us show now that Y satisfies the conditions of Section 7.3, and derive the quermass densities D_i for Y.

THEOREM 7.7. *For the RACS Y having functional (7.16), the quermass densities are given by*

$$D_i = \begin{cases} 1 - \exp\left(-d_0\right) & (i = 0) \\ -\exp\left(-d_0\right) i! \dfrac{\omega_n \omega_i}{\omega_{n-i}} \sum\limits_{s=1}^{i} \dfrac{(-1)^s}{s!} a(i, s) & (i > 0) \end{cases} \tag{7.18}$$

where

$$a(i, s) = \sum_{\substack{\nu_1+\ldots+\nu_s=i \\ 1\le\nu_k\le q}} \prod_{k=1}^{s} \frac{\omega_{n-\nu_k} d_{\nu_k}}{\omega_{\nu_k} \omega_n \nu_k!} \quad (1 \le s \le i) . \tag{7.19}$$

(The summation in $a(i, s)$ refers to *ordered* s-tuples $(\nu_1, \ldots, \nu_s)$.)

Proof. We may write Y as $\bigcup\limits_{j=1}^{\infty} C_j$, where C_j are the cylinders (or grains, in the case $q = n$) of the underlying point process. From the additivity of χ ,

$$E[\chi(Y \cap Q)] = E \sum \chi\left(Q \cap C_j\right) - E \sum_{j<k} \chi\left(Q \cap C_j \cap C_k\right) + \ldots$$

$$= \sum_{k=1}^{\infty} (-1)^{k-1} p_k E\left(N_k\right) \quad (Q \in C) \tag{7.20}$$

where N_k is the number of k-tuples of cylinders hitting Q and p_k is the expected proportion of k-tuples which intersect with Q . The expectation and summation operators in the derivation of (7.20) may be interchanged due to the fact that

$$\sum_{k=1}^{\infty} E\left|(-1)^{k-1} p_k N_k\right| \le \sum_{1}^{\infty} EN_k \le \sum E\binom{N_1}{k} < \infty , \tag{7.21}$$

as N_1 is a Poisson random variable with finite mean. From (7.17),

$$EN_1 = \lambda\mu(Q) \tag{7.22}$$

where

$$\mu(Q) = \frac{1}{\omega_n} \sum_{i=0}^{q} \binom{q}{i} W_{n-i}(Q) \overline{W_i^q} \quad \text{and} \quad \overline{W_i^q} = E\left[W_i^q\left(G^q\right)\right] . \tag{7.23}$$

Hence

$$EN_k = [\lambda\mu(Q)]^k/k! \quad (k \geq 1) \ . \tag{7.24}$$

To evaluate p_k , we use a result of Streit (1970). If k convex cylinders $C_1, \ldots, C_k$ having compact convex q-dimensional bases $G_1^q, \ldots, G_k^q$ are given independent IUR positions hitting Q , then the probability that they intersect with Q is

$E[\chi(Q \cap C_1 \cap \ldots \cap C_k)]$

$$= \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} \Psi_i \left(G_1^q, \ldots, G_k^q\right) W_{n-i}(Q) / \prod_{j=1}^{k} \mu\left(Q, G_j^q\right) \tag{7.25}$$

where

$$\Psi_i\left(G_1^q, \ldots, G_k^q\right) = \frac{i!\omega_i\omega_n}{\omega_{n-i}} \sum_{\nu_1+\ldots+\nu_k=i} \prod_{j=1}^{k} \frac{\binom{q}{\nu_j}\omega_{n-\nu_j}}{\binom{n}{\nu_j}\nu_j!\omega_{\nu_j}\omega_n} W_{\nu_j}^q\left(G_j^q\right) \tag{7.26}$$

and $\mu\left(Q, G_j^q\right)$ is given by the expression (7.23) for $\mu(Q)$ with $\overline{W}_i^q$ replaced by $W_i^q\left(G_j^q\right)$. In (7.26), Ψ_i is understood to equal zero if there is no k-tuple $(\nu_1, \ldots, \nu_k)$ satisfying the conditions $\sum \nu_j = i$, $0 \leq \nu_j \leq q$.

Conditional upon a k-tuple of cylinders of the Poisson process hitting Q , the joint distribution of their positions is independent IUR. We therefore need to evaluate

$$E\left[\Psi_i\left(G_1^q, \ldots, G_k^q\right) / \prod_{j=1}^{n} \mu\left(Q, G_j^q\right) \;\middle|\; C_j \uparrow Q, \ 1 \leq j \leq k\right] . \tag{7.27}$$

Cylinders with larger μ value are more likely to hit Q . Hence this conditional moment is found by weighting the unconditional moment according to the factor $\prod_{j=1}^{n} \mu\left(Q, G_j^q\right)$, yielding

$$\overline{\Psi}_{i,k}/[\mu(Q)]^k \tag{7.28}$$

where $\overline{\Psi}_{i,k}$ is obtained by substituting $\overline{W}^q_{\nu_j}$ for $W^q_{\nu_j}\left(G^q_j\right)$ in (7.26).

We can now substitute this expected value into (7.25) to obtain

$$p_k = \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} \overline{\Psi}_{i,k} W_{n-i}(Q)/[\mu(Q)]^k . \tag{7.29}$$

Hence, from (7.20),

$$E[\chi(Y \cap Q)] = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{\lambda^k}{k!\omega_n} \sum_{i=0}^{n} \binom{n}{i} \overline{\Psi}_{i,k} W_{n-i}(Q)$$

$$= \frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} \left[\sum_{k=1}^{\infty} (-1)^{k-1} \frac{\lambda^k}{k!} \overline{\Psi}_{i,k} \right] W_{n-i}(Q) . \tag{7.30}$$

We note that this is a continuous functional of Q, and by comparison with (7.4),

$$D_i = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{\lambda^k}{k!} \overline{\Psi}_{i,k} . \tag{7.31}$$

For $i = 0$, $\overline{\Psi}_{0,k} = \left(\overline{W}^q_0\right)^k$, and hence

$$D_0 = 1 - \exp\left(-d_0\right) , \tag{7.32}$$

as claimed. Equation (7.31) can be further simplified for $i > 0$. The product $\prod_{j=1}^{k}$ appearing in the expression for $\overline{\Psi}_{i,k}$ may be subdivided as $\prod_{\nu_j=0} \prod_{\nu_j>0}$, so that

$$\overline{\Psi}_{i,k} = \frac{i!\omega_n\omega_i}{\omega_{n-i}} \sum_{s=1}^{i} \binom{k}{s} \left(\overline{W}^q_0\right)^{k-s} \sum_{\substack{\nu_1+\ldots+\nu_s=i \\ 1\le\nu_j\le q}} \prod_{j=1}^{s} \frac{\binom{q}{\nu_j}\omega_{n-\nu_j}}{\binom{n}{\nu_j}\omega_{\nu_j}\omega_n\nu_j!} \overline{W}^q_{\nu_j} \quad (i > 0) . \tag{7.33}$$

By changing the order of summation in (7.31),

$$D_i = \sum_{s=1}^{i} \frac{i!\omega_n\omega_i(-1)^{s-1}}{\omega_{n-i}s!}\left(\sum_{k=s}^{\infty} \frac{(-1)^{k-s}\lambda^{k-s}\left(\overline{W}_0^q\right)^{k-s}}{(k-s)!}\right)a(i,\ s) \quad (i > 0)\ . \tag{7.34}$$

The theorem (7.7) follows immediately. This formula embraces some results on line segment processes by Parker and Cowan (1976) and generalizes equation (3.4) of Davy (1976).

The values of D_i are tabulated below for the practical cases of 1, 2 and 3 dimensions. Firstly putting $q = n$, we are considering processes of intervals on the line, compact convex sets in the plane, and compact convex sets in space respectively.

i \ n	1	2	3
0	$1-\exp(-\lambda\overline{L})$	$1-\exp(-\lambda\overline{A})$	$1-\exp(-\lambda\overline{V})$
1	$2\lambda\ \exp(-\lambda\overline{L})$	$\frac{1}{2}\lambda\overline{B}\ \exp(-\lambda\overline{A})$	$\frac{1}{3}\lambda\overline{S}\ \exp(-\lambda\overline{V})$
2	-	$\left(\pi\lambda-\frac{1}{4}\lambda^2\overline{B}^2\right)\exp(-\lambda\overline{A})$	$\left(\frac{2}{3}\pi\lambda\overline{M}-\pi^2\lambda^2\overline{S}^2/96\right)\exp(-\lambda\overline{V})$
3	-	-	$\left(\frac{4}{3}\pi\lambda-\frac{2}{3}\pi\lambda^2\overline{M}\overline{S}+\pi^2\lambda^3\overline{S}^3/288\right)\exp(-\lambda\overline{V})$

TABLE 7.2. Values of D_i for $q = n$

As an example, consider the case of 2 dimensions. For a compact convex set Q, we have from (7.4) and Table 7.2,

$$E[A(Q \cap Y)] = A(Q)\{1-\exp(-\lambda\overline{A})\}\ , \tag{7.35}$$

$$E[B(Q \cap Y)] = \lambda A(Q)\overline{B}\ \exp(-\lambda\overline{A}) + B(Q)\{1-\exp(-\lambda\overline{A})\}\ , \tag{7.36}$$

$$E[C(Q \cap Y)] = \lambda A(Q)\left(2\pi-\tfrac{1}{2}\lambda\overline{B}^2\right)\exp(-\lambda\overline{A}) + \lambda\overline{B}B(Q)\exp(-\lambda\overline{A}) + 2\pi\{1-\exp(-\lambda\overline{A})\}\ . \tag{7.37}$$

In the case when the grains are actually line segments of mean length $\overline{L}$, we have $\overline{A} = 0$ and $\overline{B} = 2\overline{L}$. Hence

$$E[A(Q \cap Y)] = 0\ , \tag{7.38}$$

$$E[L(Q \cap Y)] = \tfrac{1}{2}E[B(Q \cap Y)] = \lambda A(Q)\overline{L}\ , \tag{7.39}$$

$$E[C(Q \cap Y)] = \lambda A(Q)\left(2\pi-2\lambda\overline{L}^2\right) + 2\lambda\overline{L}B(Q)\ . \tag{7.40}$$

To obtain the mean number of intersections within Q, recall that the mean number of segments hitting Q is $\lambda\mu(Q) = \lambda A(Q) + \lambda B(Q)\overline{L}/\pi$. Hence

$$\text{mean number of intersections} = \lambda\mu(Q) - \frac{1}{2\pi}E[C(Q \cap Y)]$$

$$= \lambda^2\overline{L}^2 A(Q) . \qquad (7.41)$$

These results agree with those given by Parker and Cowan (1976).

Next, putting $q = n - 1$, we have a process of strips of mean width $\overline{L}$ in the plane, and cylinders with mean cross-sectional area $\overline{A}$ and perimeter $\overline{B}$ for $n = 2, 3$ respectively.

$i \backslash n$	2	3
0	$1 - \exp(-\lambda\overline{L})$	$1 - \exp(-\lambda\overline{A})$
1	$\lambda \exp(-\lambda\overline{L})$	$\frac{1}{3}\lambda\overline{B} \exp(-\lambda\overline{A})$
2	$-\lambda^2 \exp(-\lambda\overline{L})$	$\left(\frac{1}{3}\lambda\pi - \pi^2\lambda^2\overline{B}^2/96\right)\exp(-\lambda\overline{A})$
3		$\left(-\frac{1}{3}\pi\lambda^2\overline{B} + \pi^2\lambda^3\overline{B}^3/288\right)\exp(-\lambda\overline{A})$

TABLE 7.3. Values of D_i for $q = n - 1$

Finally, taking $n = 3$, $q = 1$, we have a process of slabs in space having mean thickness $\overline{L}$.

$i \backslash n$	3
0	$1 - \exp(-\lambda\overline{L})$
1	$\frac{2}{3}\lambda \exp(-\lambda\overline{L})$
2	$-\left(\pi^2\lambda^2/24\right)\exp(-\lambda\overline{L})$
3	$\left(\pi^2\lambda^3/36\right)\exp(-\lambda\overline{L})$

TABLE 7.4. Values of D_i for $q = 1$

When G^q is a.s. a point and $q = 1$, the model is a Poisson process of hyperplanes. These hyperplanes enclose a countable collection of convex polytopes. D_n is then equal to $(-1)^{n-1}\omega_n N_V$, where N_V is the number of polytopes per unit volume. From Theorem 7.7, $D_0 = 0$ and

$$D_i = \frac{\omega_n \omega_i}{\omega_{n-i}} (-1)^{i-1} \left(\frac{2\omega_{n-1}\lambda}{\pi\omega_n} \right)^i \quad (0 < i \le n) . \tag{7.42}$$

The mean polygon characteristics are given by

$$E(W_i) = (-1)^{i-1} D_i / N_V = \frac{\omega_i}{\omega_{n-i}} \left(\frac{\pi\omega_n}{2\lambda\omega_{n-1}} \right)^{n-i} \quad (0 < i \le n) \tag{7.43}$$

and

$$E(W_0) = (N_V)^{-1} = \frac{1}{\omega_n} \left(\frac{\pi\omega_n}{2\lambda\omega_{n-1}} \right)^n , \tag{7.44}$$

as has been demonstrated by Miles (1972b).

7.6. Further models

As shown in Section 7.4, random sets may be constructed from simpler ones by performing certain set operations.

Suppose that $Y_1, \ldots, Y_m$ are independent Poisson RACS with parameters d_{ij} $(0 \le i \le n,\ 1 \le j \le m)$. From (7.8) and (7.17),

$$\begin{aligned} T_{\cup Y_j}(Q) &= 1 - \prod_{j=1}^{m} \exp\left\{ -\frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} d_{ij} W_{n-i}(Q) \right\} \quad (Q \in C) \\ &= 1 - \exp\left\{ -\frac{1}{\omega_n} \sum_{i=0}^{n} \binom{n}{i} \left[\sum_{j=1}^{m} d_{ij} \right] W_{n-i}(Q) \right\} . \end{aligned} \tag{7.45}$$

The union $\bigcup_1^m Y_j$ retains the stationary, isotropic, infinitely-divisible and semi-Markov character of its constituent RACS. The quermass densities are given by (7.18) with d_i replaced by $\sum_{j=1}^{m} d_{ij}$.

Similarly, these four properties are retained when an independent a.s. compact convex RACS is added to Y_1 .

The intersection RACS $\bigcap_{j=1}^{m} Y_j$ retains the first three properties but not semi-Markovicity. From (7.9), (7.16), we obtain for $m = 2$,

$$T_{Y_1 \cap Y_2}(Q) = 1 - E\left[\exp\left\{-\lambda_2 E\left(V_q\left(G_2^q - (Q \cap Y_1)^*\right) \middle| Y_1\right)\right\}\right] . \tag{7.46}$$

The quermass densities for $\bigcap_{j=1}^{m} Y_j$ may be obtained from Corollary 7.5 and Theorem 7.7 - we omit the general expressions here as they are cumbersome. In particular,

$$D_0\left(\bigcap_{j=1}^{m} Y_j\right) = \prod_{j=1}^{m} [1-\exp(-d_{0j})] \tag{7.47}$$

and

$$D_1\left(\bigcap_{j=1}^{m} Y_j\right) = \sum_{j=1}^{m} d_{1j} D_0\left(\bigcap_{k \neq j} Y_k\right) \exp(-d_{0j}) . \tag{7.48}$$

The Poisson grain model described in the previous section consists of grains which may interpenetrate each other, making it unsuitable for many applications. The mathematical treatment of non-overlapping convex particles in E^n seems much more difficult (see Ripley (1977)).

One such model is obtained as follows. To each grain G_i^n of a Poisson grain process assign a random variable ξ_i such that the ξ_i are independently and identically distributed uniformly on the interval $[0, 1]$. Let Y be the union of all grains G_i^n such that $G_j^n \uparrow G_i^n \Rightarrow \xi_j > \xi_i$, or in other words the union of all grains which are not obscured by grains of smaller ξ value. Then Y consists a.s. of disjoint grains. The probability that a grain G^n is retained is

$$\int_0^1 \exp(-\lambda\xi\mu(G^n))\,d\xi = \frac{1-\exp(-\lambda\mu(G^n))}{\lambda\mu(G^n)} \tag{7.49}$$

where

$$\mu(G^n) = \frac{1}{\omega_n} \sum_{j=1}^{n} \binom{n}{j} \overline{W}_j W_{n-j}(G^n) . \tag{7.50}$$

For example, if the Poisson model consists of overlapping balls with radius distribution function $F(t)$, then

$$D_i = \int_0^\infty \frac{t^{n-i}\left[1-\exp\{-\lambda\omega_n\int_0^\infty(t+s)^n F(ds)\}\right]}{\int_0^\infty(t+s)^n F(ds)} F(dt) . \tag{7.51}$$

The original radius distribution is shifted towards the smaller radii, as small balls are less likely to be obscured.

Another model is obtained by letting Y be the union of all grains G_i^n such that $G_j^n \uparrow G_i^n \Rightarrow V\left(G_j^n\right) < V\left(G_i^n\right)$. Such a model is useful for spatial processes where competition occurs (see Gates (1978)). The expression for D_i is complicated in general, but in the case of Poisson balls with radius distribution F , the probability that a ball of radius t is retained is

$$\exp\left\{-\lambda\omega_n \int_t^\infty (t+s)^n F(ds)\right\} \tag{7.52}$$

and hence

$$D_i = \lambda\omega_n \int_0^\infty t^{n-i} \exp\left\{-\lambda\omega_n \int_t^\infty (t+s)^n F(ds)\right\} F(dt) . \tag{7.53}$$

In this case the new radius distribution F^* is shifted towards both the small and large radii, as the small balls are less likely to be hit at all and the large balls are less likely to be hit by still larger balls. Hence a bimodal radius distribution may be obtained. For example, if F is uniform on $[0, 1]$, $n = 2$ and $\lambda = \pi^{-1}$, the density f^* of the modified distribution is bimodal with minimum at $(1+\sqrt{8})/7$. $f^*(t)/f^*(1)$ is tabulated below.

t	0	.2	.4	.547	.6	.8	1
$f^*(t)/f^*(1)$	.717	.574	.475	.450	.454	.561	1

TABLE 7.5. Modified radius density for areal competition process

CHAPTER 8

THICK SECTIONS

8.1. Introduction

One of the important areas of application of stereology is within transmission microscopy. Thin slices of partially transparent material are placed under the microscope, and the shadow cast by the opaque phase is observed. Provided that the slice is very thin, the sterelogical formulae for planar sections are approximately true. However, in general, corrections for section thickness should be used.

Miles (1974, 1976) has studied the effects of section thickness for a model of opaque Poisson-distributed particles in E^3 . In this chapter Miles' results are extended in a number of directions. We consider the union Y of a Poisson process of convex grains or cylinders in E^n . Instead of observing a true r-section, we obtain the orthogonally projected image of all points of Y within distance t of the flat F_r . The result for grains appears in Davy (1976), unfortunately together with some errors in the coefficients tabulated for particular cases.

Before considering thick sections through random sets, we develop a thick section theory for deterministic structures analogous to the theory developed for standard sections, although certain problems arise with non-convexity. Finally, some practical data is analysed.

8.2. Deterministic structure

Suppose that, instead of observing a true r-section $Y \cap F_r$ through the feature set Y , we observe a deformed image $Y'(F_r, t)$. Y' can be described in either of two ways. Firstly, it is the orthogonal projection onto F_r of the intersection of Y with a cylinder of the form $F_r + tO_{n-r}$, O_{n-r} being orthogonal to F_r . Alternatively, Y' is the set of points in F_r whose component of distance from Y in a direction orthogonal to F_r is less than or equal to t . This is just what happens in transmission microscopy when a slab of thickness $2t$ is viewed under the microscope instead of a true section, it being assumed that Y is opaque and the remainder of the specimen transparent. For convex Y , the quermassintegrals of Y' satisfy

$$\int W_i^r(Y')dF_r = \frac{\omega_r}{\omega_{r-i}} \iint_{F_i^r \uparrow Y'} dF_i^r dF_r \Big/ \int dL_i^r$$

$$= \frac{\omega_r}{\omega_{r-i}} \int dL_{r(i)} \int_{C_i(t)\uparrow Y} dC_i(t) \Big/ \int dL_i^r \qquad (8.1)$$

where $C_i(t)$ is of the form $F_i + tO_{n-r}$, O_{n-r} being orthogonal to F_i .

The total measure of cylinders hitting Y is (Streit (1970))

$$\int_{C_i(t)\uparrow Y} dC_i(t) = \frac{\int dL_i}{\omega_n} \sum_{s=0}^{n-i} \binom{n-i}{s} W_{n-s}(Y) W_s^{n-i}(tO_{n-r})$$

$$= \frac{\int dL_i}{\omega_n} \sum_{s=r-i}^{n-i} \binom{n-r}{s+i-r} W_{n-s}(Y) \frac{\omega_s \omega_{n-r}}{\omega_{s+i-r}} t^{n-i-s} \qquad \text{(by (3.9))}$$

$$= \frac{\int dL_i \omega_{n-r}}{\omega_n} \sum_{s=0}^{n-r} \binom{n-r}{s} W_{i+s}(Y) \frac{\omega_{n-i-s}}{\omega_{n-r-s}} t^s . \qquad (8.2)$$

Hence, substituting (8.2) into (8.1),

$$\int W_i^r(Y')dF_r = \frac{\omega_r \omega_{n-r}}{\omega_n \omega_{r-i}} \int dL_r \sum_{s=0}^{n-r} \binom{n-r}{s} \frac{\omega_{n-i-s}}{\omega_{n-r-s}} W_{i+s}(Y) t^s$$

$$(0 \le i \le r < n) . \qquad (8.3)$$

Note that when $t = 0$, (8.3) reduces to (3.3).

Suppose now that Y , itself opaque, is contained in a transparent convex specimen set X . It is not clear just how we should define a random thick section through X . We could define it to be an IUR flat over all positions within distance t of X. In this case, from (8.3),

$$E\left[W_i^r(Y')\right] = \frac{\frac{\omega_r}{\omega_{r-i}} \sum_{s=0}^{n-r} \binom{n-r}{s} \frac{\omega_{n-i-s}}{\omega_{n-r-s}} W_{i+s}(Y) t^s}{\sum_{s=0}^{n-r} \binom{n-r}{s} W_{r+s}(X) t^s} \quad (0 \le i \le r < n) . \tag{8.4}$$

For example, if a thick section is taken through the 3-dimensional specimen X, then the projected image Y' has expected area A and perimeter B given by

$$EA = [V(Y)+\tfrac{1}{2}tS(Y)]/[M(X)+2t] , \tag{8.5}$$

$$EB = [\pi S(Y)/4 + 2\pi t M(Y)]/[M(X)+2t] . \tag{8.6}$$

If an IUR cylinder of radius t is taken through X, then the expected length of the projected image Y' is

$$EL = \frac{V(Y)+\frac{\pi}{4}tS(Y)+\pi t^2 M(Y)}{\frac{1}{4}S(X)+\pi t M(X)+\pi t^2} . \tag{8.7}$$

Alternatively, we could just consider those flats which actually hit X. It is convenient to suppose that Y is sufficently "internal", in the sense that any flat for which $Y'(F_r, t) \neq \emptyset$ also satisfies $X \cap F_r \neq \emptyset$. Then we obtain the simpler formula

$$E\left[W_i^r(Y')\right] = \frac{\omega_r}{\omega_{r-i}} \sum_{s=0}^{n-r} \frac{\omega_{n-i-s}}{\omega_{n-r-s}} W_{i+s}(Y) t^s / W_r(X) \quad (0 \le i \le r < n) . \tag{8.8}$$

Weighted sampling is easier to formulate in this second case. For a V_r-weighted section through X,

$$E_{V_r}\left\{\frac{W_i^r(Y')}{V_r(X \cap F_r)}\right\} = \frac{\omega_r \omega_{n-r}}{\omega_n \omega_{r-i}} \sum_{s=0}^{n-r} \binom{n-r}{s} \frac{\omega_{n-i-s} W_{i+s}(Y)}{\omega_{n-r-s} V(X)} t^s . \tag{8.9}$$

For example, putting $n = 3$, $r = 2$,

$$E_A(A_A) = V_V + \frac{t}{2} S_V \tag{8.10}$$

and

$$E_A(B_A) = \frac{\pi}{4} S_V + 2\pi t M_V . \tag{8.11}$$

Equation (8.10) has been given by Underwood (1970, p. 174) (note that t in his equation corresponds to $2t$ in ours).

The preceding results do not extend readily to non-convex Y ; for example holes of maximum diameter less than $2t$ can never be detected by thick sections. In particular they do not extend to a collection of convex particles embedded in X . This is due to the difficulty of dealing with overlap - in mathematical terms projection and intersection are not commutative operations. Therefore we shall resort to random sets to deal with particle aggregates.

8.3. Thick sections through Poisson random sets

Consider the Poisson model described in Section 7.5. Clearly, the projected image of the intersection of a think section with a Poisson model is an r-dimensional Poisson model. To find the parameters $d_i^r(t)$ of the projected process, we note that if Q is an r-dimensional compact convex set, then from (7.17),

$$\frac{1}{\omega_r}\sum_{i=0}^{r}\binom{r}{i}d_i^r(t)W_{r-i}^r(Q) = \frac{1}{\omega_n}\sum_{i=0}^{n}\binom{n}{i}d_iW_{n-i}\left(Q+tO_{n-r}\right) . \qquad (8.12)$$

But from equation (51) in Hadwiger (1957, Chapter 6),

$$W_{n-i}\left(Q+tO_{n-r}\right) = \sum_{s=\max(0,r-i)}^{\min(n-i,r)} \frac{\omega_{n-i}\omega_{n-r}}{\omega_s\omega_{n-i-s}}\binom{n-r}{i+s-r}\binom{r}{s}W_s^r(Q)t^{i+s-r}\Big/\binom{n}{i} . \qquad (8.13)$$

Hence the right-hand side of (8.12) is equal to

$$\frac{1}{\omega_n}\sum_{s=0}^{r}\binom{r}{s}W_{r-s}^r(Q)\left\{\sum_{k=0}^{n-r} d_{k+s}\binom{n-r}{k}\frac{\omega_{n-k-s}\omega_{n-r}}{\omega_{r-s}\omega_{n-k-r}}t^k\right\} . \qquad (8.14)$$

By equating coefficients of $W_{r-i}^r(Q)$,

$$d_i^r(t) = \frac{\omega_r\omega_{n-r}}{\omega_n\omega_{r-i}}\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\omega_{n-i-k}}{\omega_{n-r-k}}d_{k+i}t^k \quad (0 \le i \le r) . \qquad (8.15)$$

The projected densities $D_i^r(t)$ now follow by substituting $d_i^r(t)$ for

d_i^r in Theorem 7.7.

The 2- and 3- dimensional versions of these results are as follows:

$n = 3$, $r = 2$, $q = 3$ (thick planar section through grain process)

$$A_A(t) = 1 - \exp\{-\lambda(\overline{V}+\tfrac{1}{2}t\overline{S})\} , \tag{8.16}$$

$$B_A(t) = \lambda\left(\frac{\pi}{4}\,\overline{S} + 2\pi t\overline{M}\right)\exp\{-\lambda(\overline{V}+\tfrac{1}{2}t\overline{S})\} , \tag{8.17}$$

$$C_A(t) = \left\{2\pi\lambda(\overline{M}+2t)-\tfrac{1}{2}\pi^2\lambda^2(\tfrac{1}{4}\overline{S}+2t\overline{M})^2\right\}\exp\{-\lambda(\overline{V}+\tfrac{1}{2}t\overline{S})\} . \tag{8.18}$$

$n = 3$, $r = 2$, $q = 2$ (thick planar section through cylinder process)

$$A_A(t) = 1 - \exp\{-\lambda(\overline{A}+\tfrac{1}{2}\overline{B}t)\} , \tag{8.19}$$

$$B_A(t) = \lambda\left(\frac{\pi}{4}\,\overline{B} + \pi t\right)\exp\{-\lambda(\overline{A}+\tfrac{1}{2}\overline{B}t)\} , \tag{8.20}$$

$$C_A(t) = \left\{2\pi\lambda-\tfrac{1}{2}\pi^2\lambda^2(\tfrac{1}{4}\overline{B}+t)^2\right\}\exp\{-\lambda(\overline{A}+\tfrac{1}{2}\overline{B}t)\} . \tag{8.21}$$

$n = 3$, $r = 2$, $q = 1$ (thick planar section through slab process)

$$A_A(t) = 1 - \exp\{-\lambda(\overline{L}+t)\} , \tag{8.22}$$

$$B_A(t) = \tfrac{1}{2}\lambda\pi\,\exp\{-\lambda(\overline{L}+t)\} , \tag{8.23}$$

$$C_A(t) = -\tfrac{1}{8}\lambda^2\pi^2\exp\{-\lambda(\overline{L}+t)\} . \tag{8.24}$$

$n = 3$, $r = 1$, $q = 3$ (cylindrical probe through grain process)

$$L_L(t) = 1 - \exp\left\{-\lambda\left(\overline{V} + \frac{\pi}{4}t\overline{S} + \pi t^2\overline{M}\right)\right\} , \tag{8.25}$$

$$P_L(t) = 2\lambda\left(\tfrac{1}{4}\overline{S}+\pi t\overline{M}+\pi t^2\right)\exp\left\{-\lambda\left(\overline{V} + \frac{\pi}{4}\,t\overline{S} + \pi t^2\overline{M}\right)\right\} . \tag{8.26}$$

$n = 3$, $r = 1$, $q = 2$ (cylindrical probe through cylinder process)

$$L_L(t) = 1 - \exp\left\{-\lambda\left(\overline{A} + \frac{\pi}{4}\,t\overline{B} + \frac{\pi}{2}\,t^2\right)\right\} , \tag{8.27}$$

$$P_L(t) = \lambda(\tfrac{1}{2}\overline{B}+\pi t)\exp\left\{-\lambda\left(\overline{A} + \frac{\pi}{4}\,t\overline{B} + \frac{\pi}{2}\,t^2\right)\right\} . \tag{8.28}$$

$n = 3$, $r = 1$, $q = 1$ (cylindrical probe through slab process)

$$L_L(t) = 1 - \exp\left\{-\lambda\left(\overline{L} + \frac{\pi}{2}\,t\right)\right\} , \tag{8.29}$$

$$P_L(t) = \lambda\,\exp\left\{-\lambda\left(\overline{L} + \frac{\pi}{2}\,t\right)\right\} . \tag{8.30}$$

$n = 2$, $r = 1$, $q = 2$ (strip probe through grain process in plane)

$$L_L(t) = 1 - \exp\{-\lambda(\overline{A} + 2t\overline{B}/\pi)\} , \tag{8.31}$$

$$P_L(t) = 2\lambda(\overline{B}/\pi + 2t)\exp\{-\lambda(\overline{A} + 2t\overline{B}/\pi)\} . \tag{8.32}$$

$n = 2$, $r = 1$, $q = 1$ (strip probe through strip process in plane)

$$L_L(t) = 1 - \exp\left\{-\lambda\left(\overline{L} + \frac{4}{\pi}\,t\right)\right\} , \tag{8.33}$$

$$P_L(t) = \frac{4\lambda}{\pi}\,\exp\left\{-\lambda\left(\overline{L} + \frac{4}{\pi}\,t\right)\right\} . \tag{8.34}$$

Miles (1974, 1976) has discussed how the formulae (8.16)-(8.18) may be used to estimate parameters of the model from thick section data. It is assumed either that t and at least one of λ, $\overline{V}$, $\overline{S}$ or $\overline{M}$ is known, or else that data is available from two sections of unknown but differing thicknesses.

In some cases, however, estimation is possible on the basis of sections of a single known thickness.

For example, from (8.19)-(8.21), omitting the t dependence of A_A, B_A and C_A for simplicity of notation,

$$V_V = 1 - \left(1-A_A\right)\exp\left\{\frac{t\left(2B_A-tC_A\right)}{\pi\left(1-A_A\right)} - \frac{t^2B_A^2}{2\pi\left(1-A_A\right)^2}\right\}$$

$$= A_A - \frac{2t}{\pi}\,B_A + o(t) , \tag{8.35}$$

$$S_V = \left\{\frac{2\left(2B_A-tC_A\right)}{\pi\left(1-A_A\right)} - \frac{tB_A^2}{\pi\left(1-A_A\right)^2}\right\}\exp\left\{\frac{t\left(2B_A-tC_A\right)}{\pi\left(1-A_A\right)} - \frac{t^2B_A^2}{2\pi\left(1-A_A\right)^2}\right\}$$

$$= \frac{4}{\pi}\,B_A + \frac{t}{\pi}\left[\frac{B_A^2(8-\pi)}{\pi\left(1-A_A\right)} - 2C_A\right] + o(t) , \tag{8.36}$$

$$K_V = \left\{ \frac{C_A}{1-A_A} + \frac{\pi t B_A}{1-A_A}\left[\tfrac{1}{2}C_A + \frac{B_A^2}{4(1-A_A)}\right] - \tfrac{1}{2}t^2\left[\frac{C_A}{2} + \frac{B_A^2}{4(1-A_A)}\right]^2 \right\}$$

$$\times \exp\left\{\frac{t(2B_A - tC_A)}{\pi(1-A_A)} - \frac{t^2B_A^2}{2\pi(1-A_A)^2}\right\}$$

$$= C_A + \pi t B_A\left[\tfrac{1}{2}C_A + \frac{2B_AC_A}{\pi^2(1-A_A)} + \frac{B_A^2}{4(1-A_A)}\right] + o(t) \ . \quad (8.37)$$

In the case of slabs distributed in space, the formulae are simpler. From (8.22)-(8.24),

$$V_V = 1 - (1-A_A)\exp\left\{\frac{2tB_A}{\pi(1-A_A)}\right\}$$

$$= A_A - \frac{2t}{\pi}B_A + o(t) \ , \quad (8.38)$$

$$S_V = \frac{4}{\pi}B_A \exp\left\{\frac{2tB_A}{\pi(1-A_A)}\right\} = \frac{4}{\pi}B_A \exp\left\{-\frac{4tC_A}{\pi B_A}\right\}$$

$$= \frac{4}{\pi}B_A - \frac{16t}{\pi^2}C_A + o(t) \ , \quad (8.39)$$

$$K_V = C_A \exp\left\{\frac{2tB_A}{\pi(1-A_A)}\right\}$$

$$= C_A + 2tC_AB_A/\pi(1-A_A) + o(t) \ . \quad (8.40)$$

For a cylindrical probe through a slab process (formulae (8.29), (8.30)),

$$V_V = 1 - (1-L_L)\exp\left\{\frac{\pi t P_L}{2(1-L_L)}\right\} = L_L - \frac{\pi}{2}tP_L + o(t) \ , \quad (8.41)$$

$$S_V = 2P_L \exp\left\{\frac{\pi t P_L}{2(1-L_L)}\right\} = 2P_L + \frac{\pi t P_L^2}{1-L_L} + o(t) \ . \quad (8.42)$$

And finally, for a thick transect through a strip process in the plane,

$$A_A = 1 - (1-L_L)\exp\left\{\frac{tP_L}{1-L_L}\right\} = L_L - tP_L + o(t) \ , \quad (8.43)$$

$$B_A = \frac{\pi}{2} P_L \exp\left\{\frac{tP_L}{1-L_L}\right\} = \frac{\pi}{2} P_L + \frac{\pi}{2} t \frac{P_L^2}{1-L_L} + o(t) \ . \tag{8.44}$$

8.4. A case study

Data from sections of varying thickness taken from a 3-dimensional specimen is not readily available. However 2-dimensional data may be obtained from an image analyser as follows. We begin with an isotropic, homogeneous 2-phase image on the screen and add a linear segment of length $2t$. L_L and P_L are then measured along a line perpendicular to the added segment. This is equivalent to projecting a strip of width $2t$ onto its midline. Data may be obtained for a range of values of t , and the experiment replicated for a number of different fields.

The above experiment was carried out on a Quantimet Analyser in collaboration with Dr H. Keller of the Anatomisches Institut der Universität Bern. The slides used in the analysis were taken from a lung, 25 fields being used with strips of 13 different thicknesses (including 0) on each. A typical field of view is shown in Figure 8.1.

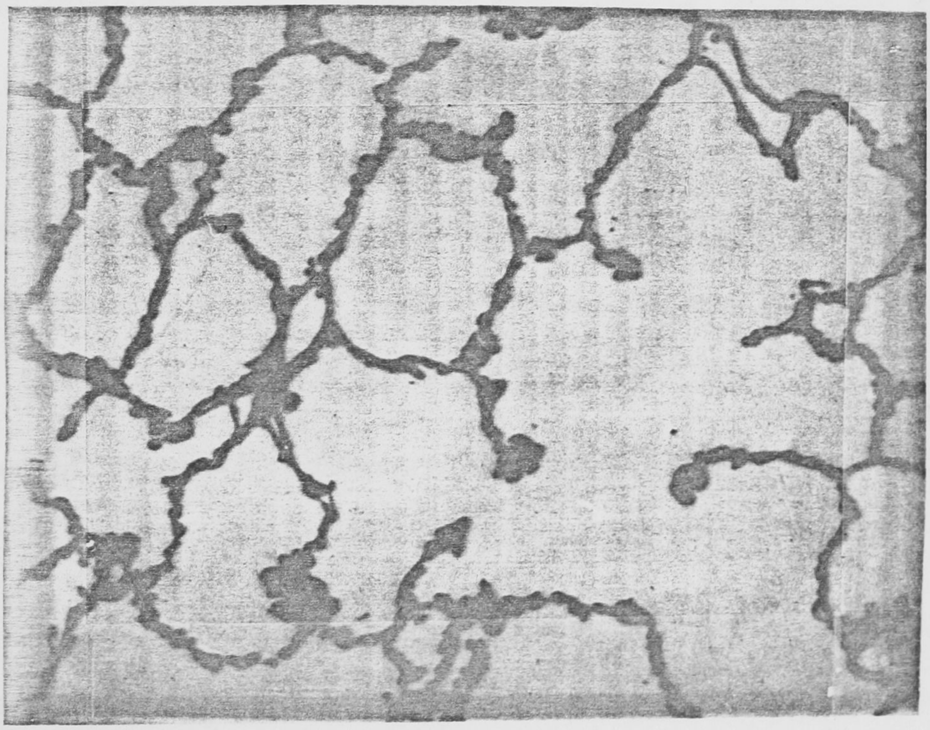


FIGURE 8.1. Section of lung used in thick section analysis

The following results were obtained (the unit of length is $8.06\ \mu$, the width corresponding to one picture point on the screen).

t	L_L	$\sigma(L_L)/5$	P_L	$\sigma(P_L)/5$
0	.1215	.0126	.235	.0216
.25	.1595	.0157	.214	.0214
.5	.1937	.0174	.205	.0213
1	.2576	.0235	.215	.0221
2	.3486	.0274	.177	.0146
3	.4397	.0351	.186	.0160
4	.5152	.0374	.185	.0175
5	.5837	.0405	.168	.0175
7.5	.7374	.0345	.145	.0175
10	.8018	.0374	.105	.0114
12.5	.8756	.0257	.086	.0115
15	.9028	.0232	.062	.0089
20	.9673	.0089	.036	.0100

TABLE 8.1. Data for 2-dimensional thick section analysis

The graph of $-\ln(1-L_L)$ vs. t (Figure 8.2) yields a straight line of gradient .1581 and intercept .1115 (least squares estimates). This is in agreement with (8.31), which predicts a straight line with gradient $2\lambda\bar{B}/\pi$ and intercept $\lambda\bar{A}$. However the trend in $P_L/(1-L_L)$ vs. t (Figure 8.3), which by (8.32) would be expected to yield a straight line with intercept .1581, is less clear. A straight line with intercept .1581 was fitted by least squares, yielding a gradient of .04277. The fitted curves are compared with the actual data in Figures 8.4 and 8.5. It can be seen that the L_L curve and the tail of the P_L curve fit very well. An examination of Figure 8.1 suggests why the agreement is so poor for small values of t on the P_L curve. Firstly, the structure does not really appear to consist of overlapping particles, as the theory requires. Secondly and more importantly, there are many indentations on the perimeter

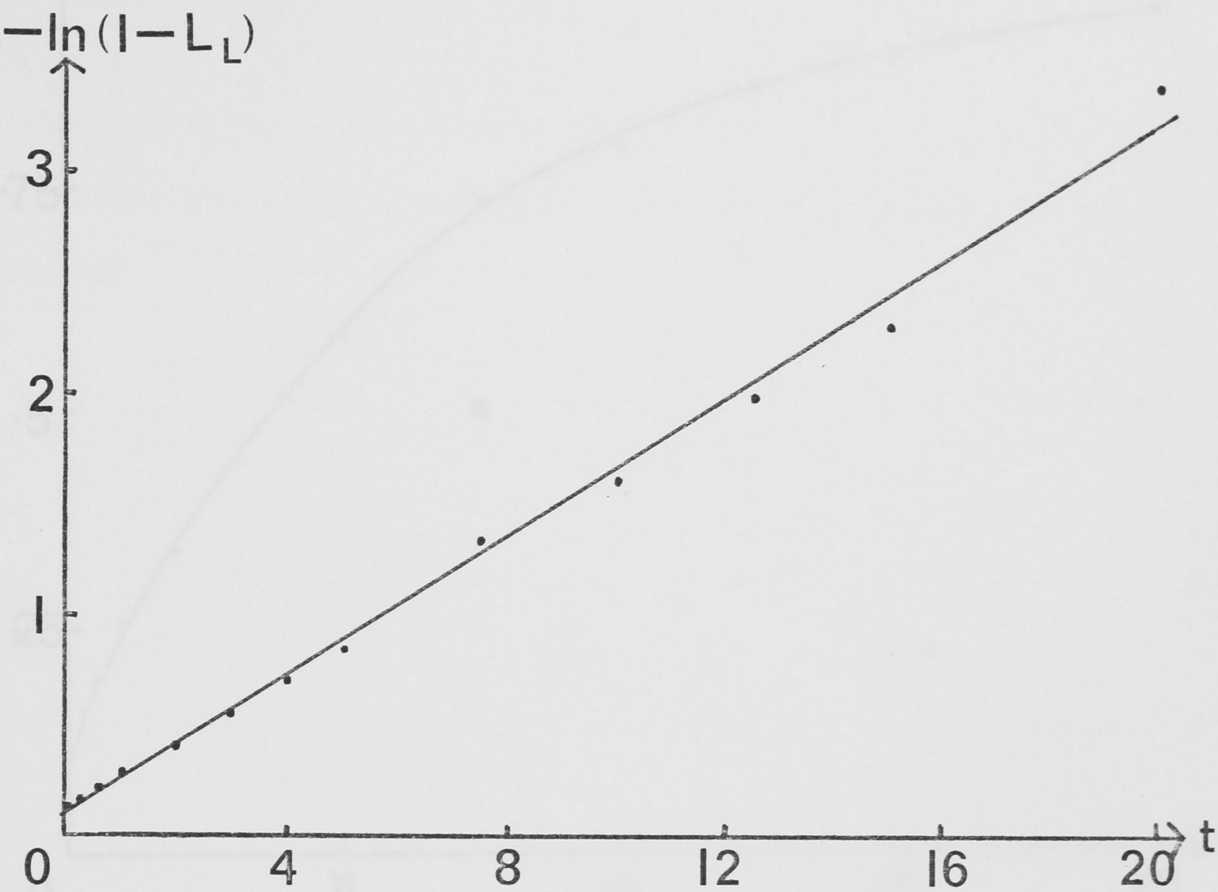


FIGURE 8.2

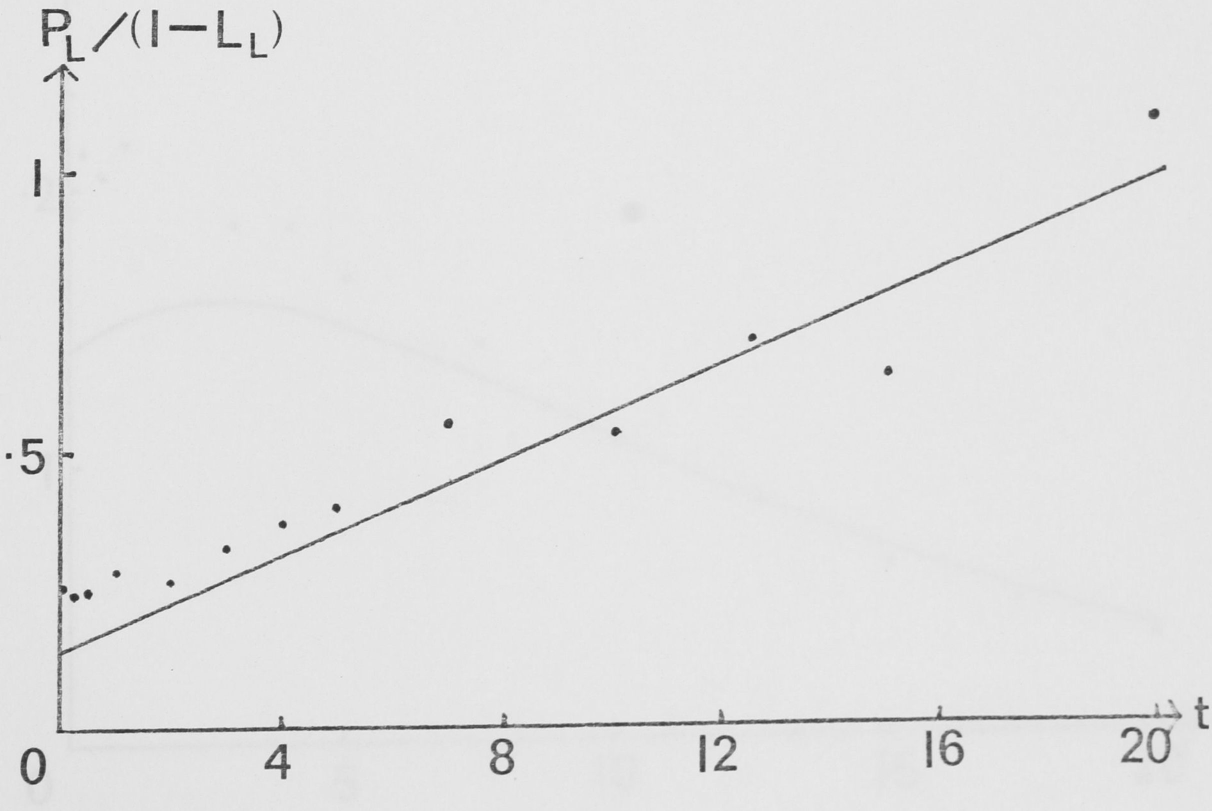


FIGURE 8.3

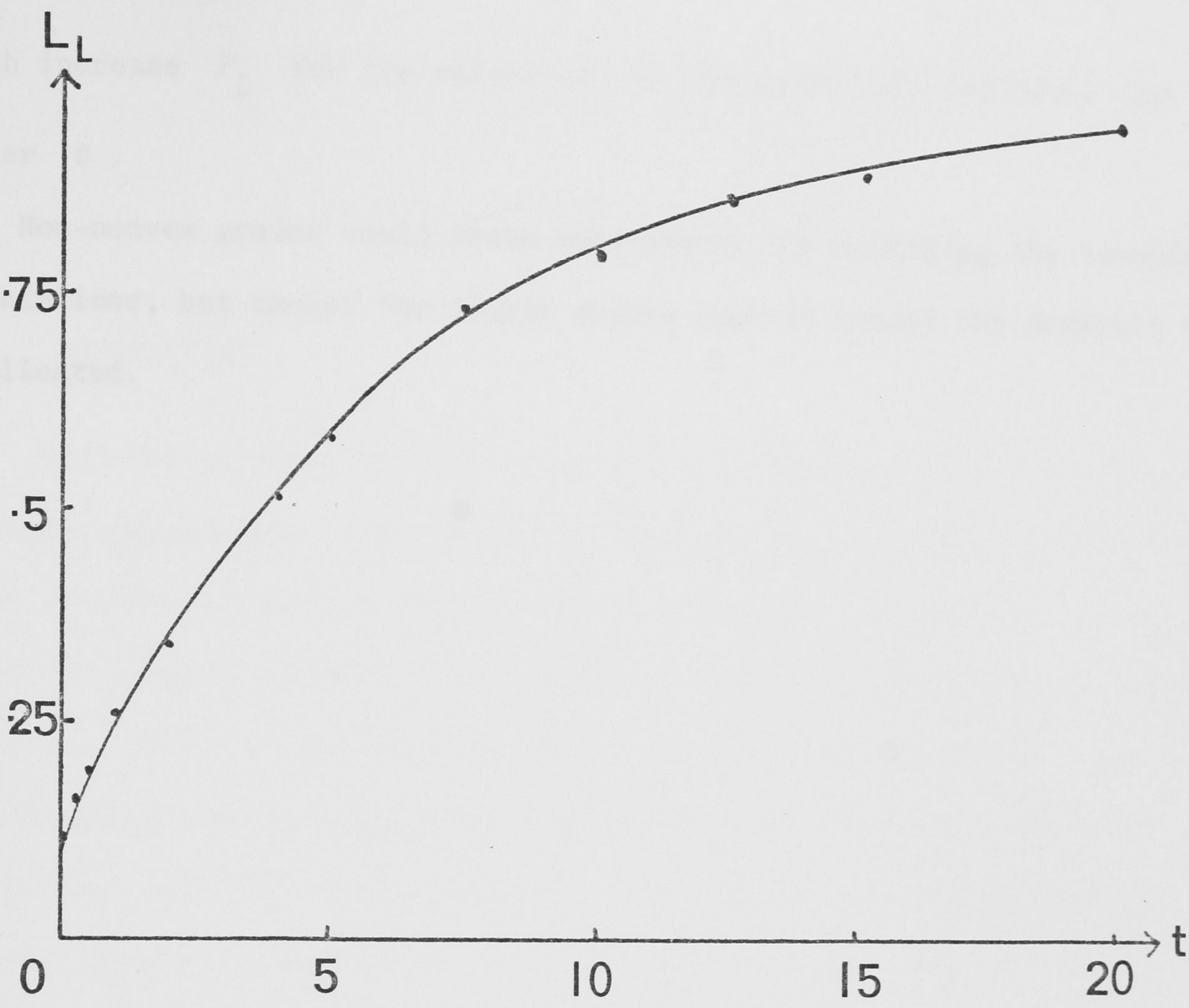


FIGURE 8.3

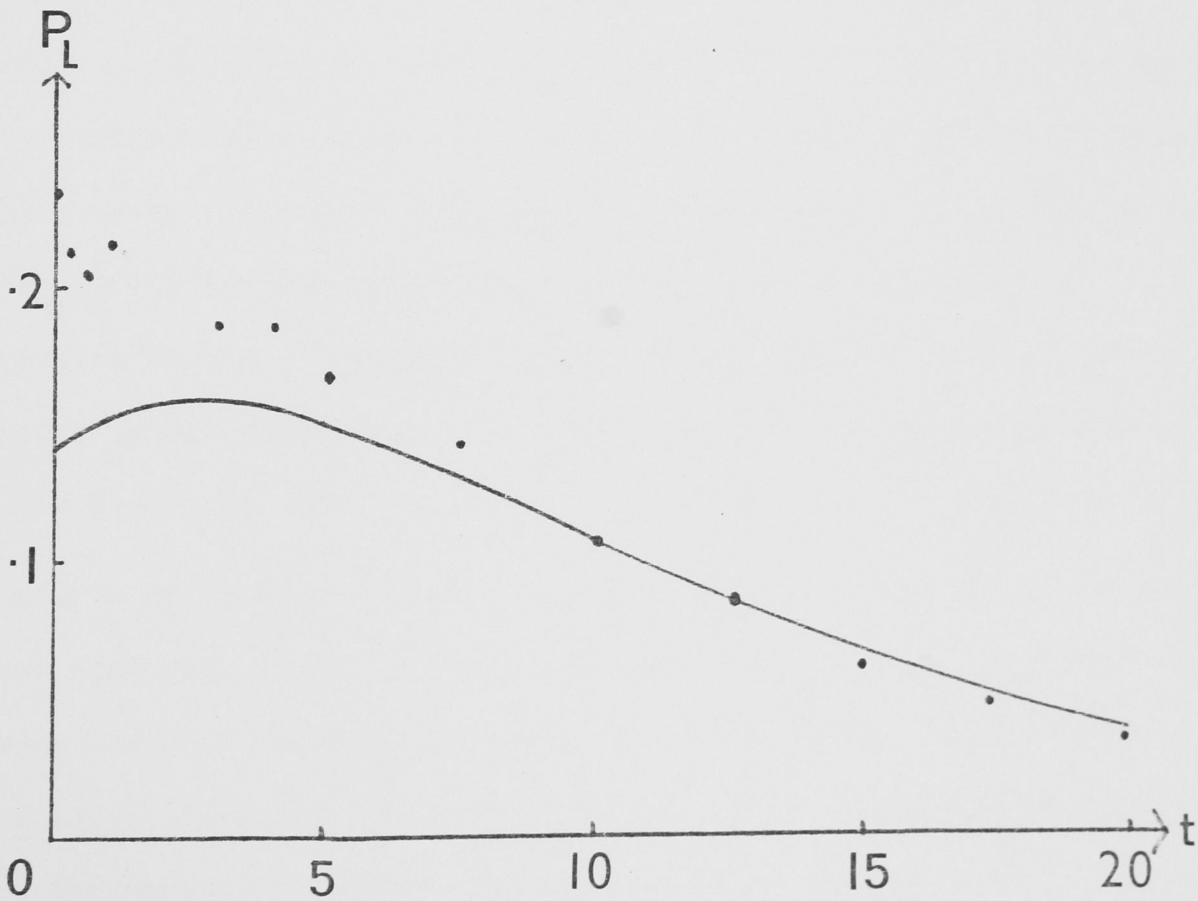


FIGURE 8.4

which increase P_L for low values of t but exert less influence for larger t .

Non-convex grains could prove more useful for modelling the boundary indentations, but except for simple shapes such as annuli the analysis is complicated.

CHAPTER 9

MULTIPLE AND MULTISTAGE SAMPLING

9.1. Introduction

Until now we have restricted ourselves to estimation on the basis of a single observation. Very rarely in practice does a stereologist arrive at an estimate after observing a single test probe. Rather, he takes a number of samples and pools his results in order to obtain a more accurate estimate. This is especially important in the anisotropic case. The traditional procedure adopted by statisticians is to take a number N of independent observations on the same random variable, which reduces the variance by the factor $1/N$.

A possible objection to this procedure is that the probes thus generated may happen to cluster together, leaving some parts of the specimen undersampled and others intensively measured. It seems more rational to space out the available probes systematically in order to get a a more representative sample. An additional advantage is that it is easier to use a ready-drawn grid of points, lines or squares, etc., rather than to go through the tedious procedure of generating a large number of independent probes. Systematic sampling, both weighted and otherwise, is discussed in Section 2, and independent sampling is treated in Section 3.

The reader may have noticed that estimators such as A_A have been referred to as if they could be measured exactly. Image analysers can perform just such tasks (at least to a very high level of precision) and are very powerful tools in stereology for this reason. Often though, especially in biological materials, phases can be distinguished only by expert interpretation rather than by grey level and A_A must in turn be estimated by stereology. Thus stereological analysis often proceeds through a number of stages, as is discussed in Section 4.

Finally, Section 5 is concerned with the problem of specimen destruction during measurement, as is the case, for example, when a 3-dimensional object is sliced by a number of planes. The ith plane must be placed relative to the aggregate of pieces remaining after the preceding sectioning, rather than through the original specimen.

9.2. Systematic sampling

The reason why we are considering systematic sampling before independent sampling, which would seem *a priori* to be the simpler procedure, is that systematic sampling can be regarded as the use of a single probe T^* of complex form rather than as the multiple use of a single probe T. See Miles and Davy (1977) for some examples of systematic probes. In the next chapter we shall consider the case of systematic orientations of lines through the origin. This section is devoted to periodic grids of parallel images of T, which may be either an r-flat $(0 \leq r < n)$ or an r-dimensional quadrat $(0 < r \leq n)$. In the case of r-flats, the intersection of the grid with its orthogonal subspace L_{n-r} is a cuboidal point grid - i.e. with a suitable choice of co-ordinates, points occur at $(k_1 a_1, \ldots, k_{n-r} a_{n-r})$, where k_i ranges over the integers $\mathbb{Z}$, and the a_i are non-zero constants. For a grid of quadrats, the reference points of the constituent probes form a cuboidal lattice in E^n. A motion of the grid is specified by

(i) a motion of a reference probe T, and

(ii) in the IUR case, a rotation of the grid about T.

The element of Haar measure is

$$dT^* = \begin{cases} dT & \text{(FUR case)} \\ dB_{n-r}^{n-r}\, dT & \text{(IUR case)} . \end{cases} \tag{9.1}$$

For each estimator $\alpha(T)$ satisfying $E[\alpha(T)] = \int_{T \uparrow X} \alpha(T)dT / \int_{T \uparrow X} dT = Z$, there corresponds a statistic $\alpha(T^*)$, namely the sum of $\alpha(T_i)$ for those members T_i of the grid which hit X. It may be shown using (9.1) that

$$E[\alpha(T^*)] = Em.Z \,, \tag{9.2}$$

where m is the number of grid members hitting X. Thus $\alpha(T^*)/Em$ is an unbiased estimator of Z. Table 9.1 lists Em for the various cases described above.

	FUR	IUR
r-flats	$(k_1 \ldots k_{n-r})^{-1} V_{n-r}(X \mid L_{n-r})$	$(k_1 \ldots k_{n-r})^{-1} M_{n-r}(X)$
r-quadrats	$(k_1 \ldots k_n)^{-1} V(X-Q)$	$(k_1 \ldots k_n)^{-1} E[V(X-Q)]$

TABLE 9.1. Em for periodic grids hitting X

Some examples are as follows:

(i) $r = 2$, $n = 3$, IUR flats (parallel planes at spacing k in E^3)

$$E[B(T^* \cap Y)] = \frac{\pi}{4k} S(Y) \,; \tag{9.3}$$

(ii) $r = 2$, $n = 2$. FUR grid of squares of side a located at points of square lattice of side $k > a$,

$$E[A(T^* \cap Y)] = \frac{a^2}{k^2} A(Y) \,. \tag{9.4}$$

The variance of $\alpha(T^*)$ may be related to that of $\alpha(T)$ via

$$\mathrm{Var}[\alpha(T^*)] = Em.\mathrm{Var}[\alpha(T)] + \sum_{t_i > 0} E[m(t_i)]\mathrm{Cov}(t_i) \,. \tag{9.5}$$

Here $m(t_i)$ is the number of pairs of probes at distance t_i, both of which hit X; $\mathrm{Cov}(t_i)$ is the covariance between $\alpha(T_i)$, $\alpha(T_j)$ for a pair of parallel probes at distance t_i both hitting X; and summation is over the countable set of distances occurring between members of the grid. $\mathrm{Cov}(t_i)$ is clearly zero for t_i exceeding the maximum diameter of X. Hasofer (1962) has evaluated the variance (9.5), together with a convenient approximation to it, in the case when T^* is a 2-dimensional point grid.

In Chapter 5, we considered ratio estimators for Z/U of the form $z(T)/u(T)$. For a grid, $z(T^*)/u(T^*)$ is also an unbiased estimator for Z/U provided that the position of T^* has been weighted with respect to $u(T^*)$. ($z(T^*)$, $u(T^*)$ are the sums of $z(T)$, $u(T)$ over all members of the grid hitting X.) The remaining problem is, how does one generate a u-weighted grid. The answer is given by

PROPOSITION 9.1. *The grid generated by choosing a u-weighted position for the reference probe T (and then giving the grid an* IR *rotation about T in the* IUR *case) has probability element*

$$u(T^*)dT^*/\int u(T^*)dT^* . \tag{9.6}$$

Proof. The joint element of T, T^* is

FUR case	IUR case
$u(T)dT/U$	$u(T)dB_{n-r}^{n-r}dT/U\int dB_{n-r}^{n-r}$. (9.7)

The marginal element of T^* is found by summing over all possible positions of T given T^*, yielding (9.6).

9.3. Independent sampling

Instead of regular grids, probes may be placed independently through the specimen. This is a more time-consuming procedure, but has the advantages that there is no danger of periodicities in the structure of Y coinciding with those of the sample, and that the variance of certain estimators may be estimated by the sample variance.

When we wish to use weighting in conjunction with independent sampling there are (at least) three different unbiased procedures. Firstly, the N probes can be weighted individually according to the factor u, and the estimator $\Lambda_1 = \frac{1}{N} \sum_{i=1}^{N} (z_i/u_i)$ used for Z/U. The variance of Λ_1 is simply

$$\mathrm{Var}(\Lambda_1) = \frac{1}{N} \mathrm{Var}_u(z/u) = \frac{1}{N} E(z^2/u) \;/\; Eu - \frac{1}{N} (Z/U)^2 . \qquad (9.8)$$

The other two methods are adapted from classical sampling theory (Cochran (1963, p. 176)). When N (> 1) IUR or FUR probes are used, it is possible to estimate the bias given in Proposition 5.2 and hence to add a correction factor to the biased estimator $\frac{1}{N} \sum_1^N z_i/u_i$. Thus

$$\Lambda_2 = \frac{1}{N} \sum_{i=1}^{N} z_i/u_i + \frac{1}{(N-1)Eu} \left[\sum_1^N z_i - \frac{1}{N} \left(\sum_1^N u_i \right) \left(\sum_1^N z_i/u_i \right) \right] \qquad (9.9)$$

is an unbiased estimator of Z/U with respect to unweighted probes (the independence is crucial, so that Λ_2 cannot be used for a grid). Denoting the correction factor in Λ_2, namely the second term in (9.9), by λ,

$$\mathrm{Var}(\Lambda_2) = \frac{1}{N} E(z/u)^2 + \frac{N-1}{N} [E(z/u)]^2 - (Z/U)^2 + E(\lambda^2) + 2E\left(\lambda \frac{z}{u}\right) . \quad (9.10)$$

Alternatively, the first probe can be u-weighted and the remainder left unweighted. By symmetry,

$$\begin{aligned} E_{u_1}\left(\sum_1^N z_i \;/\; \sum_1^N u_i \right) &= E\left(u_1 \sum z_i \;/\; \sum u_i \right) \;/\; Eu \\ &= \frac{1}{N} E\left(\sum u_i \sum z_i \;/\; \sum u_i \right) \;/\; Eu \\ &= Ez/Eu = Z/U . \end{aligned} \qquad (9.11)$$

Let the estimator within parentheses on the left-hand side of (9.11) be denoted by Λ_3. The use of such an estimator in the context of stereology was first suggested by Miles (1978). Its variance is

$$\mathrm{Var}(\Lambda_3) = E_{u_1}\left[\sum z_i \,/\, \sum u_i\right]^2 - (Z/U)^2$$

$$= \frac{1}{N} E\left[\left(\sum z_i\right)^2 \,/\, \sum u_i\right] \,/\, Eu - (Z/U)^2$$

$$= E\left[z_1^2 \,/\, \sum u_i\right] \,/\, Eu + (N-1)E\left[z_1 z_2 \,/\, \sum u_i\right] \,/\, Eu - (Z/U)^2 \,. \tag{9.12}$$

A comparison of the efficiencies of Λ_1, Λ_2 and Λ_3 seems difficult without further assumptions concerning z and u .

9.4. Multistage sampling

Stereological estimation via a number of stages can be formulated in terms of martingale theory. A series of subsamples $X \supset X_1 \supset X_2 \ldots \supset X_N$ is drawn from the specimen X . Associated with each X_i is a random variable Λ_i , which is an estimator of some property Λ_0 of X . We suppose that only Λ_N is actually observable. The underlying probability space is of the form $\mathop{\times}\limits_{j=1}^{N} \left(\Omega_j, S_j, P_j\right)$, where P_j is a probability measure on Ω_j with respect to the σ-algebra S_j . On each Ω_j is defined a conditional RACS $\tilde{X}_j$ belonging to the space $F\left(X_{j-1}\right)$ of compact subsets of X_{j-1} $\left(j > 0,\ X_0 \equiv X\right)$ in such a way that $\tilde{X}_j\left(\omega_j, X_{j-1}\right)$ is measurable both with respect to Ω_j and to $F\left(X_{j-1}\right)$. The unconditional RACS X_j can be represented as

$$X_j\left(\omega_1, \ldots, \omega_N\right) = \tilde{X}_j\left(\omega_j, \tilde{X}_{j-1}\left(\omega_{j-1}, \ldots, \tilde{X}_1\left(\omega_1, X\right) \ldots\right)\right) \tag{9.13}$$

and Λ_j as

$$\Lambda_j\left(\omega_1, \ldots, \omega_N\right) = \tilde{\Lambda}_j\left(X_j\left(\omega_1, \ldots, \omega_N\right)\right) , \tag{9.14}$$

where $\tilde{\Lambda}_j$ is a measurable, real-valued function. This formulation is represented diagramatically in Figure 9.1:

$$
\begin{tikzcd}
\Omega_1 \arrow[d, "\tilde{X}_1"'] & \Omega_2 \arrow[d, "\tilde{X}_2"'] & \cdots & \Omega_N \arrow[d, "\tilde{X}_N"'] \\
F(X_0) \arrow[d, "\tilde{\Lambda}_1"'] \arrow[ur] & F(X_1) \arrow[d, "\tilde{\Lambda}_2"'] \arrow[ur] & \cdots \arrow[ur] & F(X_{N-1}) \arrow[d, "\tilde{\Lambda}_N"'] \\
\mathbb{R} & \mathbb{R} & \cdots & \mathbb{R}
\end{tikzcd}
$$

FIGURE 9.1. Martingale formulation of multistage sampling

Λ_i is assumed to be a martingale with respect to the increasing sequence of σ-algebras $G_i = \overset{i}{\underset{j=1}{\times}} S_j \times \overset{N}{\underset{j=i+1}{\times}} \{\emptyset, \Omega_j\}$. Thus

$$E\left(\Lambda_i \mid G_{i-1}\right) = \Lambda_{i-1} \quad (1 < i \le N) \ , \tag{9.15}$$

and

$$\mathrm{Var}\left(\Lambda_i\right) = \mathrm{Var}\left(\Lambda_1\right) + \sum_{j=2}^{i} E\left[\mathrm{Var}\left(\Lambda_j \mid G_{j-1}\right)\right] \quad (1 < i \le N) \ . \tag{9.16}$$

Anderssen (1975) and Nicholson (1976) have given decompositions of variance similar to (9.15) in the case $i = 2$.

Notice that the overall variance increases as additional stages are added to the experiment. As in classical sampling theory, we may consider the problem of optimal cost allocation to multistage sampling schemes. Suppose that the total cost J of the experiment is equal to the sum $\sum J_i$, J_i being the cost of the ith stage. Thus

$$EJ = EJ_1 + E\left[\sum_{i=2}^{N} E\left(J_i \mid G_{i-1}\right)\right] \ . \tag{9.17}$$

Typically, J_i is non-random and $E\left[\mathrm{Var}\left(\Lambda_i \mid G_{i-1}\right)\right]$ is a function of the costs $J_1, \ldots, J_i$, say $v_i\left(J_1, \ldots, J_i\right)$ $\left(\mathrm{Var}\left(\Lambda_1\right) \equiv v_1\left(J_1\right)\right)$. The problem of optimal allocation is to minimize $\sum v_i\left(J_1, \ldots, J_i\right)$ subject to the condition $\sum J_i \leq J$, J being fixed. The techniques of Cochran (1963, p. 97) may be readily adapted to a stereological setting.

As an example, consider the case where N_1 independent A-weighted planes are taken through the 3-dimensional body X , and then N_2 independent UR points are taken on each planar section. X_1 consists of the N_1 planar sections and X_2 consists of the N_1N_2 random points. $\Lambda_0 = V_V$, $\Lambda_1 = \sum (A_A)_j/N_1$ and $\Lambda_2 = \sum (P_P)_i/N_1$. Now,

$$\mathrm{Var}(\Lambda_2|G_1) = \sum_{i=1}^{N_1} (A_A)_i(1-(A_A)_i)/N_1^2N_2 . \tag{9.18}$$

Hence

$$\begin{aligned} E[\mathrm{Var}(\Lambda_2|G_1)] &= \left[E(A_A)-E(A_A)^2\right]/N_1N_2 \\ &= V_V(1-V_V)/N_1N_2 - \mathrm{Var}(A_A)/N_1N_2 \\ &= k_2/N_1N_2 - k_1/N_1N_2 \text{, say.} \end{aligned} \tag{9.19}$$

Also,

$$\mathrm{Var}(\Lambda_1) = k_1/N_1 . \tag{9.20}$$

Suppose that the cost of a planar section is j_1 , and that of a random point is j_2 . Then $J_1 = N_1j_1$, $J_2 = N_1N_2j_2$, and

$$\mathrm{Var}(\Lambda_2) = j_1k_1/J_1 + j_2(k_2-k_1)/J_2 . \tag{9.21}$$

Using the method of Lagrange multipliers to minimize $\mathrm{Var}(\Lambda_2)$ subject to $J_1 + J_2 \leq J$, we obtain that the optimal solution is to take integer values of N_1, N_2 close to

$$N_1 = J/j_1\left[1 + \sqrt{j_1k_1/j_2(k_2-k_1)}\right] , \quad N_2 = j_1\sqrt{j_1k_1}/j_2\sqrt{j_2(k_2-k_1)} .$$

9.5. Fragmentation

Suppose that a convex specimen X is sectioned by a random hyperplane. There are now two pieces, and as "glueing" these back together is not generally practicable, further sectioning must proceed by choosing one of these pieces and generating a random hyperplane through it. At the next stage, one of the three resulting pieces is sectioned, and so on. For IUR sectioning, the distribution of particle size after the Nth stage depends upon the actual shape of X in a rather complex manner. However if V_{n-1}-weighted hyperplanes are used, and the pieces for sectioning are chosen with probability proportional to their volumes, the problem can be reduced to independent UR points on a line segment.

PROPOSITION 9.2. *The expected empiric distribution function* H_1 *of volume of the two pieces resulting from a* V_{n-1}*-weighted hyperplane cutting a convex body* X *of volume* V *is uniform on* $[0, V]$.

Proof. As demonstrated in Chapter 2, a V_{n-1}-weighted hyperplane may be constructed by choosing independently an IR $(n-1)$-subspace L_{n-1} and a UR point P of X , and then placing a hyperplane with orientation L_{n-1} through P . For $0 < x \leq V/2$, $H_1(x)$ is equal to $\frac{1}{2}$ times the probability that the smaller piece is less than or equal to x. For any orientation of L_{n-1} , the set of positions of P for which the above event occurs has volume $2x$, and hence

$$H_1(x) = \tfrac{1}{2}.2x/V = x/V \quad (x \leq V/2) . \tag{9.22}$$

As the sum of the two pieces is V , the empiric distribution function and hence also H_1 must satisfy the property $H_1(V-x) + H_1(x) = 1$. Therefore H_1 is uniform on $[0, V]$.

The above sectioning technique results in the uniform random division of the volume of the piece chosen at each stage. The expected distribution $H_N(x)$ of volume at the Nth stage is the same as the distribution of lengths obtained when the interval $[0, V]$ is divided into $(N+1)$ parts by N independent UR points. $1 - H_N(x)$ is just equal to the probability that none of these points fall into $[0, x]$, i.e.

$$H_N(x) = 1 - (1 - x/V)^N \quad (0 \le x \le V) . \tag{9.23}$$

As $N \to \infty$, H_N tends to the Weibull distribution $1 - \exp(-x/EX)^N$.

The distribution H_N^* of the volume of the Nth piece X_N chosen for sectioning is obtained by weighting H_{N-1} according to x, i.e.

$$H_N^*(x) = 1 - (1 - x/V)^{N-1}\{1 + (N-1)x/V\} \quad (N > 1) , \tag{9.24}$$

the first moment being $2V/(N+1)$. (For $N = 1$, $X_N \equiv X$.)

Let us consider estimation of the volume fraction $V(Y)/V$, $Y \subset X$. If $V_{n-1}(Y \cap X_i \cap (F_{n-1})_i)/V_{n-1}(X_i \cap (F_{n-1})_i)$ is the ratio estimator of $V(Y \cap X_i)/V(X_i)$ corresponding to the ith section, then

$$E_{V,V_{n-1}}\left[\frac{V_{n-1}(Y \cap X_i \cap F_{n-1})}{V_{n-1}(X_i \cap (F_{n-1})_i)}\right] = E_V\left[\frac{V(Y \cap X_i)}{V(X_i)}\right]$$

$$= V(Y)/V(X) = V_V . \tag{9.25}$$

The obvious estimator to use is some type of average of our N ratios. The usual average is unacceptable as it places equal importance on all pieces X_i, when in fact for large i these pieces become very small and less representative of the total specimen. Instead, some weighted average of the type

$$\sum_{i=1}^{N} k_i V_{n-1}(Y \cap X_i \cap (F_{n-1})_i)/V_{n-1}(X_i \cap (F_{n-1})_i) \tag{9.26}$$

should be used, where $k_1, \ldots, k_N$ is a decreasing set of constants which sum to 1 . A natural choice is to take k_i proportional to $EV(X_i)$, i.e. to $1/(i+1)$.

CHAPTER 10

MONTE CARLO TECHNIQUES IN STEREOLOGY

10.1. Introduction

In Chapter 8 we saw that knowledge of the type of random set under examination enables estimation from thick section data. In the deterministic case as well, prior or subjective information concerning the feature set Y may be incorporated into the sampling and estimation procedures. In this chapter we show how certain techniques from the Monte Carlo estimation of integrals (see Hammersley and Handscomb (1964), Chapter 5) can be adapted for the purpose of improving stereological estimates.

Recall that in Chapter 4 we considered estimation of the form

$$E\alpha = \int_{T\uparrow X} \left[\frac{\alpha(T)}{\int_{T\uparrow X} dT}\right] dT = Z = \int g(T)dT \text{ , say,} \tag{10.1}$$

where T is an IUR or FUR probe hitting the specimen set X. In other words, α is the classical Monte Carlo estimate of the integral Z of $g(T)$ over the positions of T hitting X. Hasofer (1962) has remarked upon this Monte Carlo interpretation in the case when T is a point lattice.

The next three sections investigate the use of importance sampling, control variates and antithetic variates within the context of stereology. In the final section, a simulation is presented.

10.2. Importance sampling

Roughly speaking, the idea behind importance sampling is to sample more intensively the "important" parts of the specimen. An *importance* *factor* ω is a strictly positive, integrable function such that

$$\int_{T\uparrow X} \omega(T)dT = 1 \ . \tag{10.2}$$

From (10.1), we obtain

$$\int_{T\uparrow X} \frac{g(T)}{\omega(T)} [\omega(T)dT] = Z \ . \tag{10.3}$$

In other words, the ratio g/ω is an unbiased estimator of Z with respect to an ω-weighted probe (one which is generated according to the probability density ω).

The ratio estimators of Chapter 5 are examples of importance sampling. In the case of a 3-dimensional specimen, we may put $T = F_2$, $g(T) = A\{Y \cap F_2\}M(X)/2\pi$, $Z = V(Y)$, and $\omega(T) = A\{X \cap F_2\}M(X)/2\pi V(X)$. The estimator g/ω of $V(Y)$ is just $V(X)A\{Y \cap F_2\}/A\{X \cap F_2\} = V(X)A_A$.

The variance of the importance estimator is related to the variance of the conventional estimator α by

$$\mathrm{Var}_{\omega}(g/\omega) = \mathrm{Var}(\alpha) - \mathrm{Cov}\{\alpha^2/\omega, \omega\} \ . \tag{10.4}$$

This variance is equal to zero in the extreme case $\omega(T) = g(T)/Z$. In practice, one should choose ω in such a way as to achieve a large positive value of the covariance term appearing in (10.4).

To illustrate the concept of importance sampling, suppose that we want to estimate $V(Y)$ for a cylindrical specimen X with known structural trend - specifically, it is known that the proportion of volume occupied by Y is approximately proportional to $\omega(x)$, where x is the height above the base and $\int_0^h \omega(x)dx = 1$. We may proceed by taking a planar section $\underline{F}_2$ parallel to the base with distance from the base distributed according to density $\omega(x)$. The appropriate estimator is $A\{Y \cap \underline{F}_2\}/\omega$.

10.3. Control variates

There is a second standard Monte Carlo method which makes use of positive correlation. We suppose that there exists a *control variate* $\lambda(T)$ such that $E\lambda = \Lambda$ is known. We may write $Z = \Lambda + (Z-\Lambda)$, where the difference $Z - \Lambda$ can be estimated by stereological methods. The variance of such a *difference* estimator is

$$\mathrm{Var}(\alpha-\lambda+\Lambda) = \mathrm{Var}(\alpha) + \mathrm{Var}(\lambda) - 2\,\mathrm{Cov}(\alpha, \lambda) , \tag{10.5}$$

which is equal to zero in the extreme case $\lambda = \alpha$.

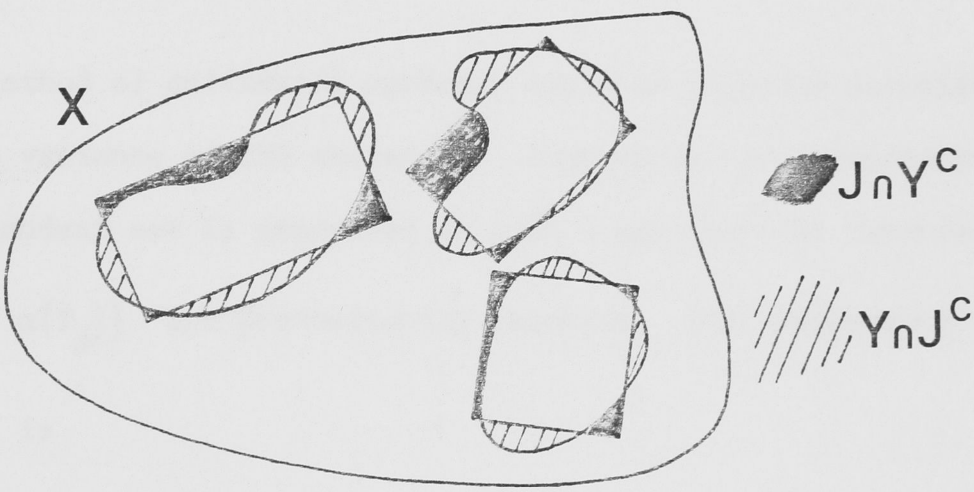


FIGURE 10.1. Use of control variate in form of approximating rectangles

Figure 10.1 illustrates the use of a control variate in estimating the area $A(Y)$. Y is approximated by a collection J of rectangles whose combined area Λ may easily be determined. We estimate $A(Y) - \Lambda$ by subtracting a stereological estimate of $A[J \cap Y^C]$ from one of $A[Y \cap J^C]$. The efficiency of such a method depends upon the accuracy of the initial approximation. If, for example, T is a UR point of X , the variance of the difference estimator of $A(Y)$ is

$$A^2(X)\left[\varepsilon_1-\varepsilon_1^2+\varepsilon_2-\varepsilon_2^2+2\varepsilon_1\varepsilon_2\right] . \tag{10.6}$$

where $\varepsilon_1 = A[J \cap Y^C]/A(X)$, $\varepsilon_2 = A[Y \cap J^C]/A(X)$.

As a further illustration, consider the example of the previous section. Suppose that the approximate constant of proportionality between volume fraction and $\omega(x)$ is known, or in other words an approximate value Λ is known for $V(Y)$. To implement difference estimation, an FUR section parallel to the base of the cylinder is generated, and $V(Y)$ is estimated by

$$(1-\omega h)\Lambda + hA(Y \cap \underline{F}_2) \ . \tag{10.7}$$

10.4. Antithetic variates

The method of *antithetic variates* exploits negative correlation to reduce the variance of the estimator. Instead of N independent probes T_i , a dependent set is generated in such a way that the covariances $\mathrm{Cov}(\alpha(T_i), \alpha(T_j))$ are predominantly negative. The variance of the average $\sum \alpha(T_i)/N$ is

$$\mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N} \alpha(T_i)\right) = \sum_{i=1}^{N} \mathrm{Var}[\alpha(T_i)]/N^2 + 2\sum_{i<j}\sum \mathrm{Cov}[\alpha(T_i), \alpha(T_j)]/N^2 \ , \tag{10.10}$$

which can be smaller than the corresponding variance for the independent case (the first term on the right-hand side).

As an example of the use of antithetic variates, consider the estimation of the length of a curvilinear feature Y contained in a planar specimen via linear transects (see Moran (1966a)). We have already seen in Chapter 5 that for an L-weighted secant of X , $E_L(\pi P_L/2) = B_A$. For N independent L-weighted secants (either independent or otherwise), the average $\pi \sum (P_L)_i/2N$ is an unbiased estimator of B_A . We now consider how to reduce the variance of this estimator by introducing dependences between the secants. One rather natural procedure is even spacing of the orientations, i.e. a uniform random angle in $[0, \pi)$ is chosen, and then N independent L-weighted FUR secants $T_1, \ldots, T_N$ are generated having orientations θ , $\left(\theta + \frac{\pi}{N}\right) \bmod \pi, \ldots, \left(\theta + \frac{(N-1)}{N}\pi\right) \bmod \pi$. Each secant then has a marginal isotropic L-weighted distribution. Now,

$$E_{L_i, L_j}[(P_L)_i, (P_L)_j] = \frac{\frac{\pi^2}{4} \iiint P(T_i \cap Y) P(T_j \cap Y) dT_i dT_j d\theta}{\iiint L(T_i \cap X) L(T_j \cap X) dT_i dT_j d\theta}$$

$$= \frac{\pi}{4} (B_A)^2 \int_0^\pi \beta\left(\theta + \frac{i-1}{N}\pi\right) \beta\left(\theta + \frac{j-1}{N}\pi\right) d\theta \quad (10.11)$$

where $\beta(\theta)$ is the total projection of Y in the direction θ divided by the length B of Y .

Hence the reduction (not necessarily positive) in variance achieved by systematically spacing the secant orientations is

$$\frac{N-1}{N} (B_A)^2 \left(1 - \frac{\pi}{4(N-1)} \sum_{i=1}^{N-1} \int_0^\pi \beta(\theta)\beta\left(\theta + \frac{i\pi}{N}\right) d\theta\right) . \quad (10.12)$$

In the extreme case when Y consists of parallel line segments, say w.l.o.g. with orientation 0 , $\beta(\theta) = |\sin\theta|$ and (9.12) becomes

$$\frac{N-1}{N} (B_A)^2 \left(1 - \frac{\pi}{4(N-1)} \sum_{i=1}^{N-1} \left\{\sin\left(\frac{i\pi}{N}\right) + \frac{\pi(N-2i)}{2N} \cos\left(\frac{i\pi}{N}\right)\right\}\right) . \quad (10.13)$$

For $N = 2, 3, 4$, the reduction in variance is $.1073(B_A)^2$, $.0761(B_A)^2$ and $.0579(B_A)^2$ respectively.

10.5. A simulation

In Chapter 5, ratio estimators used in conjunction with weighted sampling schemes were advocated for two main reasons:

(i) for *inhomogeneous* specimens, there can be considerable bias in using the fundamental formulae without weighted sectioning, and

(ii) for sufficiently *homogeneous* specimens, the mean square error is reduced by weighted sectioning.

However the criteria for optimality of weighted sampling are not easy to check in practice. A simulation was therefore carried out in two dimensions to compare the three types of estimators IUR non-ratio, IUR ratio and weighted ratio. The set analysed is shown in Figure 10.2. It comprises two specimen sets $X_1 \supset X_2$ containing a common feature set Y. Observe that Y is fairly homogeneous with respect to X_2 but not with respect to X_1.

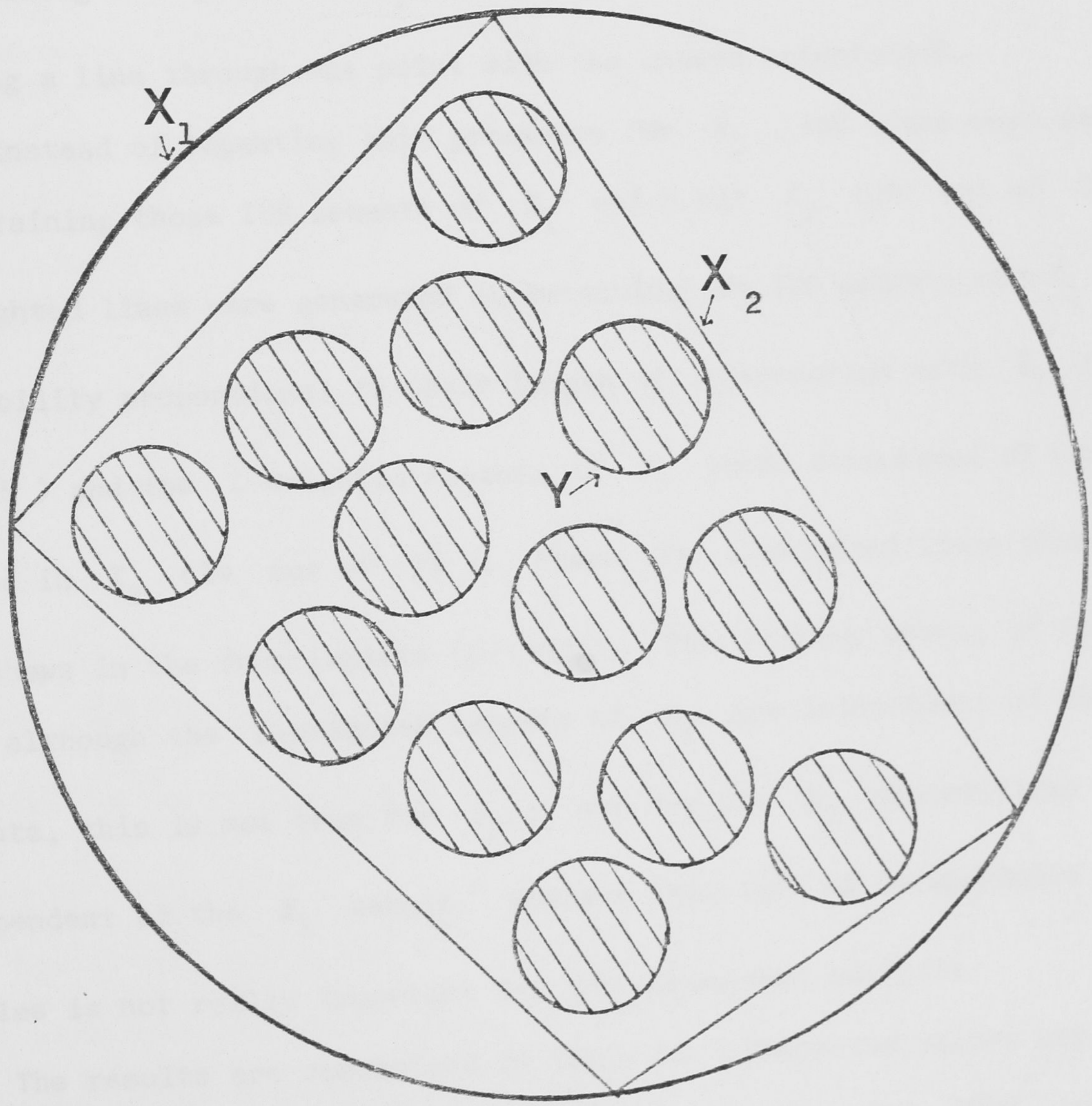


FIGURE 10.2. Planar feature set Y contained in two specimen sets X_1 and X_2.

The true areas and perimeters are as follows (the unit of length is an inch):

	X_1	X_2	Y
A	20.43	11.65	5.59
B	16.02	14.07	30.22

25 independent IUR lines and 25 independent L-weighted lines were drawn through X_1. Each IUR line was generated by using random number tables to choose an IR angle in $[0, 2\pi)$, and independent UR distance in $[0, 2.05]$, and then drawing a line at this distance from the origin and perpendicular to the direction chosen. Each L-weighted line was generated by choosing a UR point of X_1 and an independent IR angle in $[0, \pi)$, and drawing a line through the point with the chosen orientation.

Instead of repeating this procedure for X_2, IUR lines were selected by retaining those IUR secants of X_1 which hit X_2 (24 out of 25). L-weighted lines were generated by retaining the IUR secants of X_1 with probability proportional to their length of intersection with X_2 (10 out of 25) and the L-weighted secants of X_2 whose associated UR points landed in X_2 (14 out of 25). These 24 L-weighted lines through X_2 are shown in the frontispiece (p. (vi)). This economy means, of course, that although the L-weighted secants of X_1 are independent of the IUR secants, this is not true for X_2; nor are the X_2 observations independent of the X_1 sample. However this lack of independence between samples is not really important for the subsequent analysis.

The results are summarized in Table 10.3 (expected values are given in brackets), and the corresponding estimates of $A(Y)$ and $B(Y)$ appear in Table 10.4.

	X_1		X_2	
	IUR	L-weighted	IUR	L-weighted
$\overline{L} = \overline{L(X \cap F_1)}$	4.201 (4.006)	4.388 (4.329)	2.717 (2.602)	3.451
$\overline{l} = \overline{L(Y \cap F_1)}$	1.262 (1.096)	1.418	1.315 (1.248)	1.738
$\overline{P} = \overline{P(\partial Y \cap F_1)}$	4.080 (3.773)	4.800	4.250 (4.296)	5.667
$\overline{L_L}$	.2769	.3021 (.2737)	.4249	.4936 (.4798)
$\overline{P_L}$	.8970	.9867 (.9418)	1.381	1.6149 (1.651)

TABLE 10.3. Results of random line simulation associated with Figure 10.2

Estimator	$\widehat{A(Y)}$	X_1	X_2	$\widehat{B(Y)}$	X_1	X_2
IUR non-ratio (unbiased)	$M(X)\overline{l}$	6.436	5.888	$\frac{\pi}{2} M(X)\overline{P}$	32.69	29.90
IUR ratio (biased)	$A(X)\overline{L}_L$	5.657	4.952	$\frac{\pi}{2} A(X)\overline{P}_L$	28.78	25.27
Corrected IUR ratio (unbiased)	$A(X)\overline{L}_L + \frac{25}{24} M(X)\left(\overline{l}-\overline{L}.\overline{L}_L\right)$	6.182	5.670	$\frac{\pi}{2} \left\{A(X)\overline{P}_L + \frac{25}{24} M(X)\left(\overline{P}-\overline{L}.\overline{P}_L\right)\right\}$	31.39	28.93
L-weighted ratio (unbiased)	$A(X)\overline{L}_L$	6.171	5.752	$\frac{\pi}{2} A(X)\overline{P}_L$	31.66	29.56

TABLE 10.4. Estimates of area and perimeter of Y

From a cursory inspection of Table 10.4, it appears that the estimators for X_2 usually fare better than the corresponding ones for X_1 . Also, the bias in the second row is only apparent for X_2 ; for X_1 it is masked by sample error.

Let us now investigate the mean suqare errors for single observations which were discussed in Chapter 5. One approach would be to use sample variances of the above data, but this would disregard known parameters and would only provide one variance estimate for each different estimator. The following alternative procedure was adopted. From the theory presented in Chapter 5,

$$[M(X)]^2 E(l^2) - [A(Y)]^2 = \mathrm{MSE}[M(X)l] \ , \qquad (10.14)$$

$$[A(X)]^2 E(L_L)^2 + [A(Y)]^2 - 2A(Y)A(X)E(L_L) = \mathrm{MSE}[A(X)L_L] \ , \qquad (10.15)$$

$$A(X)M(X)E(l^2/L) - [A(Y)]^2 = \mathrm{MSE}_L[A(X)L_L] \ . \qquad (10.16)$$

Hence by calculating the sample averages of l^2, l^2/L, L_L and $(L_L)^2$ for the IUR observations, the three mean square errors on the left-hand sides of (10.14)-(10.16) may be estimated.

Similarly,

$$A(X)M(X)E_L(l^2/L) - [A(Y)]^2 = \mathrm{MSE}[M(X)l] \ , \qquad (10.17)$$

$$[A(X)]^3 E_L(l^2/L^3)/M(X) + [A(Y)]^2 - 2A(Y)[A(X)]^2 E_L(l/L^2)/M(X)$$

$$= \mathrm{MSE}[A(X)L_L] \ , \qquad (10.18)$$

$$[A(X)]^2 E_L(L_L)^2 - [A(Y)]^2 = \mathrm{MSE}_L[A(X)L_L] \ . \qquad (10.19)$$

Thus in the case of L-weighted observations, sample averages of l^2/L, $(L_L)^2$, l^2/L^3 and l/L^2 can be used to estimate the three types of mean square error.

Analogous expressions to (10.14)-(10.19) exist for perimeter estimators.

The above technique was used to provide two estimates of mean square error for each of the three types of single observation estimators of area and perimeter. Table 10.5 lists these MSE estimates, the first in each cell being based on the IUR sample and the second on the L-weighted sample.

The order of increasing MSE within each column is the same for the two types of estimate. Note that differences in MSE are very marked, so that choice of the correct estimator and sampling scheme is important.

	Area		Perimeter	
	X_1	X_2	X_1	X_2
IUR non-ratio (unbiased)	30.32 26.85	18.21 19.79	616.6 653.8	315.7 414.5
IUR ratio (biased)	13.20 13.08	6.935 10.57	305.7 332.6	175.5 271.1
L-weighted ratio (unbiased)	21.27 17.98	6.485 7.797	397.1 410.8	38.37 72.77

TABLE 10.5. Estimates of MSE for single observation estimators

For the inhomogeneous case (X_1) , the preferred estimator for both area and perimeter is actually biased IUR ratio, followed by L-weighted ratio and then IUR non-ratio.

For the homogeneous case (X_2) , the L-weighted ratio estimator is best,followed by IUR ratio and lastly IUR non-ratio.

However these orders only apply to estimators based on single observations. As the sample size N increases, the MSE of the unbiased estimators decreases as $1/N$, but the MSE of the biased estimator tends to the square of the bias (see Miles (1978)). For sufficiently large sample sizes therefore, L-weighted ratio estimation is optimal for both area and perimeter estimation in both X_1 and X_2 .

BIBLIOGRAPHY

(Page numbers of occurrences in the text appear in square brackets at the end of each reference.)

Anderssen, R.S. (1975) Recent results in volume-fraction analysis. Unpublished manuscript, Australian National University [106].

Anderssen, R.S. & Jakeman, A.J. (1975) Abel type integral equations in stereology. II. Computational methods of solution and the random spheres approximation. *J. of Microscopy* 105, 135-153. [3]

Baddeley, A. (1977) A fourth note on recent research in geometrical probability. *Advances Appl. Prob.* (to appear). [5]

Blaschke, W. (1949) *Vorlesungen über Integralgeometrie*. Chelsea, New York. [5, 66]

Bonnesen, T. & Fenchel, W. (1948) *Theorie der konvoxen Körper*. Chelsea, New York. [26]

Buffon, G.L.L. Comte de (1777) Essai d'arithmetique morale. *Supplément à l'Histoire Naturelle* 4. [5]

Cahn, J.W. (1967) The significance of average mean curvature and its determination by quantitative metallography. *Trans. A.I.M.E.* 239, 610-616. [2, 69]

Chen, Chang-Shing (1973) On the kinematic formula of square of mean curvature vector. *Indiana Univ. Mathematics J.* 22, 1163-1169. [67]

Cochran, W.G. (1963) *Sampling Techniques* (2nd ed.). John Wiley & Sons, New York. [,].

Crofton, M.W. (1885) Probability. Encyclopaedia Britannica, 9th ed., 19, 768-788. [5]

Davy, P.J. (1976) Projected thick sections through multi-dimensional particle aggregates. *J. Appl. Prob.* 13, 714-722. [(ii), 83, 88]

Davy, P.J. (1977) Aspects of random set theory. *Advances Appl. Prob.* (to appear). [(ii), 4]

Davy, P.J. & Miles, R.E. (1977) Sampling theory for opaque spatial specimens. *J.R. Statist. Soc. B* 39, 56-65. [(ii), 48]

DeHoff, R.T. (1967) The quantitative estimation of mean surface curvature. *Trans. A.I.M.E.* 239, 617-621. [2, 69]

DeHoff, R.T. & Rhines, F.N. (eds.) (1968) *Quantitative Microscopy*. McGraw-Hill, New York. [3]

Delesse, M.A. (1847) Procédé mécanique pour determiner la composition des roches. *C.R. Acad. Sci. (Paris)* 25, 544-545. [1]

Elias, H. (ed.) (1967) *Stereology* (Proceedings of the Second International Congress for Stereology, Chicago, 1967). Springer-Verlag, New York. [1, 2]

Federer, H. (1969) *Geometric Measure Theory*. Springer-Verlag, Berlin-Heidelberg. [29, 30]

Flanders, H. (1963) *Differential Forms with Applications to the Physical Sciences*. Academic Press, New York. [6]

Gates, D.J. (1978) Bimodality in even-aged plant monocultures. *J. Theor. Biol.* (to appear). [87]

Gel'fand, I.M., Graev, M.I. & Vilenkin, N.Y. (1966) *Generalized Functions, Vol. 5: Integral Geometry and Representation Theory.* Academic Press, New York. [31]

Giger, H. (1970) Grundleichungen der Stereologie I. *Metrika* 16, 43-57. [47]

Giger, H. (1975) Rasterbare Texturen. *Zeitschr. für ang. Math. Phys.* 26, 7-22. [72]

Glagoleff, A.A. (1933) On the geometrical methods of quantitative mineralogic analysis of rocks. *Tr. Inst. Econ. Min. and Metal.*, Moscow, 59, 475. [1]

Hadwiger, H. (1957) Vorlesungen über Inhalt, Oberfläche und Isoperimetrie. Springer, Berlin. [5, 16, 20, 21, 22, 39, 73, 91]

Hadwiger, H. (1959) Normale Körper im euklidischen Raum und ihre topologischen und metrischen Eigenschaften. *Math. Zeitschr.* 71, 124-140. [22]

Hadwiger, H. & Schneider, R. (1971) Vektorielle Integralgeometrie. *Elem. der Math.* 26, 49-72. [22]

Hammersley, J.M. & Handscomb, D.C. (1965) *Monte Carlo Methods.* Methuen, London and John Wiley & Sons, New York. [111]

Hasofer, A.M. (1963) On the reliability of the point counter method in petrography. *Aust. J. Appl. Sci.* 14, 168-179. [103, 111]

Haug, H. & Elias, H. (eds.) (1963) *Proceedings of the First International Congress for Stereology.* Congress Print, Vienna. [2]

Hicks, N.J. (1965) *Notes on Differential Geometry.* Van Nostrand, Princeton-New York. [25]

Jakeman, A.J. & Anderssen, R.S. (1975) Abel type integral equations in stereology. I. General discussion. *J. of Microscopy* 105, 121-133. [3]

Kendall, D.G. (1974) Foundations of a theory of random sets. In *Stochastic Geometry*, ed. E.F. Harding & D.G. Kendall, John Wiley & Sons, London, 322-376. [70]

Kendall, M.G. (1961) *A Course in the Geometry of n Dimensions.* Hafner, New York. [35]

Kendall, M.G. & Moran, P.A.P. (1963) *Geometrical Probability.* Hafner, New York. [5, 8]

Little, D.V. (1974) A third note on recent research in geometrical probability. *Advances Appl. Prob.* 6, 103-130. [5]

Mandelbrot, B.B. (1977) *Fractals: Form, Chance and Dimension.* W.H. Freeman, San Francisco. [29]

Matheron, G. (1975) *Random Sets and Integral Geometry.* Wiley, New York. [5, 6, 22, 27, 27, 70, 71, 76, 78, 79]

Mayhew, T.M. & Cruz-Orive, L.-M. (1974) Caveat on the use of the Delesse principle of areal analysis for estimating component volume densities. *J. Microscopy* 102, 195-207. [47]

Meyer, P.A. (1966) *Probabilités et Potential.* Hermann, Paris. [71]

Miles, R.E. (1971a) Isotropic random simplices. *Advances Appl. Prob.* 3, 353-382. [9, 16, 35, 36]

Miles, R.E. (1971b) Poisson flats and the complementary theorem. *Advances Appl. Prob.* 3, 1-43. [78]

Miles, R.E. (1972a) Multidimensional perspectives on stereology. *J. Microscopy* 95, 181-196. [35, 47]

Miles, R.E. (1972b) The random division of space. *Suppl. Advances Appl. Prob.*, 243-266. [85]

Miles, R.E. (1974) The estimation of aggregate and overall characteristics from thick sections by transmission microscopy. *Electron Microscopy 1974*, Vol. II, ed. J.V. Sanders & D.J. Goodchild, Australian Acad. Sci., Canberra, 6-7. [88]

Miles, R.E. (1975) Direct derivations of certain surface integral formulae for the mean projections of a convex set. *Advances Appl. Prob.* 7, 818-829. [26]

Miles, R.E. (1976) On estimating aggregate and overall characteristics from thick sections by transmission microscopy. *J. Microscopy* 107, 227-234. [47, 88, 93]

Miles, R.E. (1978) The importance of proper model specification in stereology. Proceedings of the Buffon Bicentenary Symposium, Paris 1977, ed. R.E. Miles & J. Serra. Springer Lecture Notes in Biomathematics. [120]

Miles, R.E. & Davy, P.J. (1976a) Probabilistic foundations of stereology. Proceedings of the Fourth International Congress for Stereology, U.S. National Bureau of Standards Special Publication No. 431, 443-445. [(ii), 48]

Miles, R.E. & Davy, P.J. (1976b) Precise and general conditions for the validity of a comprehensive set of stereological fundamental formulae. *J. Microscopy* 107, 211-226. [(ii), 48]

Miles, R.E. & Davy, P.J. (1977) On the choice of quadrats in stereology. *J. Microscopy* 110, 27-44. [(ii),101]

Miles, R.E. & Davy, P.J. (1978) Particle number or density can be stereologically estimated by wedge sections. *J. Microscopy* 113. [(ii), 59]

Moran, P.A.P. (1966a) Measuring the length of a curve. *Biometrika* 53, 359-364. [114]

Moran, P.A.P. (1966b) A note on recent research in geometrical probability. *J. Appl. Prob.* 3, 453-463. [5]

Moran, P.A.P. (1969) A second note on recent research in geometrical probability. *Advances Appl. Prob.* 1, 73-89. [5]

Moran, P.A.P. (1972) The probabilistic basis of stereology. *Suppl. Advances Appl. Prob.*, 69-91. [3]

Nachbin, L. (1965) *The Haar Integral.* Van Nostrand, Princeton. [8, 9]

Nicholson, W.L. (1970) Estimation of linear properties of particle size distributions. *Biometrika* 57, 273-298. [3]

Nicholson, W.L. (1976) Statistics of measurements. Proceedings of the Fourth International Congress for Stereology, U.S. National Bureau of Standards Special Publication No. 431, 461-464. [106]

Parker, P. & Cowan, R. (1976) Some properties of line segment processes. *J. Appl. Prob.* 13, 96-107. [83, 84]

Ripley, B.D. (1977) Modelling spatial patterns. *J. R. Statist. Soc. B* 39, 172-192. [86]

Rosiwal, A. (1898) Über geometrische Gesteinsanalysen. Ein einfacher Weg zur ziffernmassigen Feststellung des Quantitätsverhältnisses der Mineralbestandteile gemengter Gesteine. *Verh. K.K. Geol. Reichsanst.* 5-6, 143-175. [1]

Santaló, L.A. (1955) Sur la mesure des espaces linéaires qui coupent un corps convexe et problèmes qui s'y rattachent. Collque sur les questions de réalité en géométrie, Liège. Georges Thone, Liège; Masson et Cie, Paris, 177-190. [10]

Santaló, L.A. (1976) *Integral Geometry and Geometrical Probability*. Addison-Wesley, Massachusetts. [5, 6, 36, 40]

Schneider, R. (1971) On Steiner points of convex bodies. *Israel J. Math.* 9, 241-249. [23]

Serra, J. (1969) *Introduction à la Morphologie Mathématique* Fascicule 3, Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, édité par l'École Nationale Superieure des Mines de Paris. [47]

Streit, F. (1970) On multiple integral geometric integrals and their applications to probability theory. *Canad. J. Math.* 22, 151-163. [89]

Struik, D.J. (1961) *Lectures on Classical Differential Geometry*. Addison-Wesley, Massachusetts. [61]

Todhunter, I. & Leathem, J.G. (1963) *Spherical Trigonometry*. Macmillan, London. [60]

Tomkeieff, S.T. (1945) Linear intercepts, areas and volumes. *Nature* 155, 24. [2]

Underwood, E.E. (1970) *Quantitative Stereology*. Addison-Wesley, Massachusetts. [3, 91]

Underwood, E.E., de Wit, R. & Moore, G.A. (eds.) (1976) *Proceedings of the Fourth International Congress for Stereology*, Gaithersburg, 1975. National Bureau of Standards (U.S.A.) Special Publication 431. [3]

Weibel, E.R. (1973) Stereological techniques for electron microscopic morphometry. In *Principles and Techniques of Electron Microscopy*, ed. M.A. Hayat, Van Nostrand-Reinhold, New York, 237-296. [2]

Weibel, E.R., Meek, G., Ralph, B., Echlin, P. & Ross, R. (eds.) (1971) *Stereology* 3. Proceedings of the Third International Congress for Stereology, Berne, 1971. Blackwell, Oxford. [2]

Wicksell, S.D. (1925) The corpuscle problem, Pt. I: A mathematical study of a biometric problem. *Biometrika* 17, 84-99. [3, 59]

Wicksell, S.D. (1926) The corpuscle problem. Pt. II. *Biometrika* 18, 151-172. [3, 59]

GLOSSARY OF NOTATION

The following list is restricted mostly to symbols which retain the same meaning throughout the thesis, and includes the page numbers of first appearance. Further symbols make casual appearances, but should cause no confusion as they are defined in the sections in which they are used. Due to the scarcity of Greek and Roman letters, some symbols have been used with different meangins in different contexts. Also, subscripted or superscripted letters usually have entirely different meanings to unadorned letters. Standard stereological nomenclature has been adhered to as closely as possible.

Symbol	Page	Meaning
A	35	area
A_A	3	areal fraction
B	35	perimeter or boundary length
B_r, B_r^n	16	r-frame in E^n
C	56	total curvature of planar curve, i.e. K_1^2
$\mathcal{C}$	19	set of all compact convex sets of E^n
$\mathbf{c}$	42	centroid
Cov	48	covariance
C_i, $C_i(t)$	80, 89	cylinder (of radius t)
D_i, D_i^n	74	ith quermass density
d_i, d_i^n	79	parameter of Poisson model
E^n	8	n-dimensional Euclidean space
E	12	expectation
E_u	50	expectation with respect to u-weighting
F_r, F_r^n	11	r-flat of E^n
$\underline{F}_r$	11	r-flat with fixed orientation
$F_{r(q)}$	11	r-flat containing fixed q-flat
FUR	12	fixed-orientation uniform random
f	31	real-valued function on E^n
$\mathbf{f}$	31	vector-valued function $E^n \to E^n$
$\mathbf{f}_r$	31	projection of $\mathbf{f}$ onto F_r

Symbol	Page	Meaning
F	70	space of closed subsets of E^n
G	26	integral of Gaussian curvature, i.e. K_2^3
G^q	78	grain $(q = n)$ or cylinder cross-section $(q < n)$
H	56	mean areal projection, i.e. M_2^3
h_i	26, 27	ith symmetric (generalized) function of principal curvatures
I	56	indicator function
IR	10	isotropic random
IUR	12	isotropic uniform random
I^m	30	m-dimensional integralgeometric measure
K	35	integral of mean curvature, i.e. K_1^3
K_i, K_i^m	26	integral of ith symmetric function of curvatures over hypersurface
K	21	class of all finite unions of members of C
k	65	curvature vector
k_N, k_g	65	normal curvature vector, geodesic curvature vector
L	20, 56	length
L_r, L_r^n	9	r-dimensional subspace of E^n
$L_{r(q)}$	9	r-subspace containing fixed q-subspace
M	16, 56	mean caliper diameter, i.e. M_1^2 or M_1^3
M_r, M_r^n	12	mean r-projection of compact set in E^n
MSE	51	mean square error
N	64	number of particles number of observations number of sampling stages
N	25	unit normal vector
n	8	dimension of Euclidean space
O, O_n	20	unit n-ball
P	45, 56	number of intersections or points
p_i	23	ith location functional
p_Y	40	probability density of the distance between two independent UR points of Y
Q	15	quadrat

Symbol	Page	Meaning
Q_y	15	translate of quadrat by y
$\mathbf{q}_i$	22	ith vectorial integral geometric functional
q	78	dimension of compact cylinder base for Poisson RACS
R	67	integral of squared mean curvature
RACS	71	random closed set
S, S_n	16	surface area (of hypersurface in E^n)
s	25	arc length
T	49	probe (either flat or quadrat)
T_Y	71	hitting functional of RACS
T^*	101	grid probe
$\mathbf{t}$	25	unit tangent vector
t	88	radius of thick section
UR	17	uniform random
u	50	weighting factor
U	50	$\int_{T\uparrow X} u(T)dT$
V, V_n	8	n-dimensional volume, i.e. W_0^n
V_V	3	volume fraction
Var	37	variance
W_i, W_i^n	16,19, 21	ith quermassintegral
X	8	specimen set
$\mathbf{x}_i$	11	point in E^i
Y	19 71	feature set random closed set
z	50	function of probe
Z	50	$\int_{T\uparrow X} z(T)dT$
α, α_r	34, 52	non-ratio estimator (based on r-flat)
$\boldsymbol{\alpha}$, $\boldsymbol{\alpha}_r$	41	vector-valued non-ratio estimator (based on r-flat)
Γ	9	gamma function
δ	35	distance between two flats
∂	23	boundary of set
κ	25	curvature

Symbol	Page	Meaning
κ_i	25	ith principal curvature
θ	35	angle between two flats or subspaces
σ_i	9	surface area of unit sphere in E^i $(= i\omega_i)$
ω_i	17	volume of unit ball in E^i
ω	112	importance factor
χ	21	Euler-Poincaré characteristic
$\uparrow$	11	hits
∇f	32	gradient
$\nabla . f$	32	divergence
$\nabla \times f$	32	rotation
$X\|L_r$	11	orthogonal projection of X onto L_r
$X \pm Y$	15	$\{x \pm y \mid x \in X, y \in Y\}$